THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY

CARL MURCHISON

VOLUME XLI

1 9 3 2

Published quarterly by the
CLARK UNIVERSITY PRESS
Worcester, Massachusetts
U S. A

TABLE OF CONTENTS

M BATALLA AND F DAVIS
A Life-Size Alley Maze for Children . . . 235-239

LAURETTA BENDER
Gestalt Principles in the Sidewalk Drawings and Games of Children 192-210

HAROLD BENJAMIN
Age and Sex Differences in the Toy-Preferences of Young Children 417-429

M BURSTEIN AND E. B HURLOCK
The Imaginary Playmate A Questionnaire Study 380-392

HERBERT A CARROLL AND FRANK H FINCH
Gifted Children as High-School Leaders . 476-481

W D COMMINS
A Note on the Learning of Elevated Mazes by Rats . 481-482

HERBERT S CONRAD
The Personal Equation in Ratings I An Experimental Determination . . 267-293

JOHN W COX
A Critical Note on Some Recent Research on Mechanical Ability 228-234

F DAVIS AND M BATALLA
A Life-Size Alley Maze for Children 235-239

WAYNE DENNIS AND R H HENNEMAN
The Non-random Character of Initial Maze Behavior—Erratum 463

WAYNE DENNIS, R H HENNEMAN, AND W B. VARNER
A Further Study of the Initial Maze Behavior of Rats . 358-368

WAYNE DENNIS AND J M PORTER, JR
Isolated Action of Compound Stimuli in a Locomotor Habit of Rats . 127-135

DOROTHY DUNN AND HAROLD E JONES
The Configural Factor in Children's Learning . 3-15

FRANK H FINCH AND HERBERT A CARROLL
Gifted Children as High-School Leaders 476-481

H F HARLOW
Social Facilitation of Feeding in the Albino Rat . 211-221

H F HARLOW
Food Preferences of the Albino Rat . 430-438

R H HENNEMAN AND WAYNE DENNIS
The Non-random Character of Initial Maze Behavior—Erratum 463

R H HENNEMAN, WAYNE DENNIS, AND W B VARNER
A Further Study of the Initial Maze Behavior of Rats . 358 368

JOSEPHINE ROHRS HILGARD
Learning and Maturation in Preschool Children 36-56

WALTER S HUNTER
The Effect of Inactivity Produced by Cold upon Learning and Retention in the Cockroach, *Blatella germanica* . 253-266

E B HURLOCK AND M BURSTEIN
The Imaginary Playmate, A Questionnaire Study 380-392

HAROLD E JONES AND DOROTHY DUNN
The Configural Factor in Children's Learning 3-15

ROSA KATZ
The Inventive Genius of the Child . 369-379

HELEN LOIS KOCH AND HELEN STREIT
A Study in Rating Technique with Special Reference to Activity in Preschool Children 330-357

ALICE M LEAHY
A Study of Certain Selective Factors Influencing Prediction of the Mental Status of Adopted Children, or Adopted Children in Nature-Nurture Research . . . 294-329

JOHN LEVY
A Clinical Study of the Application of Mental Hygiene to the Treatment of Children's School Problems 439-461

JULIUS B. MALLER AND JOSEPH ZUBIN
The Effect of Motivation upon Intelligence Test Scores 136-151

R H MASURE
Learning of a Petz's Conure 463-476

FRED MCKINNEY
Identical Sensory Elements versus Functional Equivalents in Visual Transfer Reactions . 483-489

JOSEPH W NAGGE
Regarding the Law of Parsimony 492-494

HARRIET E. O'SHEA
"Habits" of Three Successive Nursery-School Groups and Some Relations between Traits . . 167-191

J. M PORTER, JR, AND WAYNE DENNIS
Isolated Action of Compound Stimuli in a Locomotor Habit of Rats . . . 127-135

AUDREY M SHUEY
Some Experiments with Kittens on the Simple Alternation Problem . . . 393-405

RUTH STAPLES
Some Factors Influencing the Afternoon Sleep of Young Children . . . 222-228

HELEN STREIT AND HELEN LOIS KOCH
A Study in Rating Technique with Special Reference to Activity in Preschool Children . 330-357

RUTH UPDEGRAFF
The Determination of a Reliable Intelligence Quotient for the Young Child . 152-166

W B VARNER, WAYNE DENNIS, AND R H HENNEMAN
A Further Study of the Initial Maze Behavior of Rats 358-368

C J WARDEN
The Relative Strength of the Primary Drives in the White Rat 16-35

LUCIEN H. WARNER
The Association Span of the White Rat 57-90

LUCIEN H WARNER
An Experimental Search for the "Conditioned Response" 91-115

BETH L. WELLMAN
Some New Bases for Interpretation of the IQ . 116-126

BETH L WELLMAN
Some New Bases for Interpretation of the IQ—Erratum 462

JOSEPH G YOSHIOKA
Learning versus Skill in Rats . 406-416

JOSEPH G. YOSHIOKA
A Further Note on Size Preference of Albino Rats 489-492

JOSEPH ZUBIN AND JULIUS B MALLER
The Effect of Motivation upon Intelligence Test Scores 136-151

BOOKS

CARL MURCHISON [ED]
A Handbook of Child Psychology (Barbara S Burks) 240-248

S ZUCKERMAN
The Social Life of Monkeys and Apes (O L Tinklepaugh) 495-497

A HANDBOOK OF CHILD PSYCHOLOGY

By J E. ANDERSON, C. BUHLER, A FREUD, A. GESELL, F L GOODENOUGH, L S HOLLINGWORTH, S ISAACS, H E JONES, M C JONES, V JONES, C W KIMMINS, H KLÜVER, K LEWIN, H. MARSHALL, D MCCARTHY, M MEAD, J. PETERSON, J PIAGET, R PINTNER, L M. TERMAN, B L WELLMAN, AND H T. WOOLLEY *Edited by* CARL MURCHISON

$5 00 postpaid

A HISTORY OF PSYCHOLOGY IN AUTOBIOGRAPHY: VOLUME I

By J M BALDWIN, M W CALKINS, E CLAPARÈDE, R DODGE, P JANET, J. JASTROW, F. KIESOW, W MCDOUGALL, C. E SEASHORE, C. SPEARMAN, W STERN, C. STUMPF, H. C WARREN, T. ZIEHEN, AND H. ZWAARDEMAKER. *Edited by* CARL MURCHISON

$5.00 postpaid

A HISTORY OF PSYCHOLOGY IN AUTOBIOGRAPHY. VOLUME II

By B. BOURDON, J DREVER, K. DUNLAP, G. C FERRARI, S I FRANZ, K. GROOS, G. HEYMANS, H HOFFDING, C H JUDD, C L. MORGAN, W. B PILLSBURY, L M. TERMAN, M. F WASHBURN, R S WOODWORTH, AND R M. YERKES *Edited by* CARL MURCHISON.

$5 00 postpaid

THE FOUNDATIONS OF EXPERIMENTAL PSYCHOLOGY

By H. BANISTER, P. BARD, W B CANNON, W J CROZIER, A. FORBES, S I FRANZ, F. N. FREEMAN, A GESELL, H HARTRIDGE, S HECHT, J Q. HOLSOPPLE, W S. HUNTER, T. L. KELLEY, C LANDIS, K S LASHLEY, M. A MAY, T H MORGAN, J P. NAFE, G. H. PARKER, R PINTNER, E SHEN, L. T TROLAND, AND C. WISSLER. *Edited by* CARL MURCHISON

$6.00 postpaid

THE PSYCHOLOGICAL REGISTER· VOLUME II (1929)

Edited by CARL MURCHISON, *in cooperation with* F. C BARTLETT, S. BLACHOWSKI, K BUHLER, S DE SANCTIS, T. G. HEGGE, M MATSUMOTO, H PIERON, AND A. L. SCHNIERMANN

$6.00 postpaid

THE PSYCHOLOGICAL REGISTER. VOLUME III (1932)

Edited by CARL MURCHISON, in cooperation with A F AALL, F C. BARTLETT, S. BLACHOWSKI, T. BORCAS, K BUHLER, R. BUJAS, E CLAPARÈDE, P DAHLE, S DE SANCTIS, J. DREVER, J. ENDARA, H. HALVERSEN, T A. HUNTER, G A JAEDERHOLM, S S. KASANDJIEV, D. KATZ, K. LEWIN, H T LOVELL, A. R. LURIA, M. MATSUMOTO, A. E. MICHOTTE, L. E MIRA, J C MONTANER, E MOUCHET, H. PIÉRON, V. F PROTOPOPOW, C RADULESCU-MOTRU, K RAMUL, P. RANSCHBURG, F. M. J. A ROELS, H DE BRITO BELFORD ROXO, E J RUBIN, H. W RUIN, F SANDER, A. L. SCHNIERMANN, M ŞEKIP, N N. SEN-GUPTA, F SERACKY, E SHEN, C. SPEARMAN, D. USNADZE, AND R. W WILCOCKS; L HARDEN, *Assistant Editor.*

$10 00 postpaid

CLARK UNIVERSITY PRESS
Worcester, Massachusetts
U S A.

volume
bers $4.00

QUARTERLY
Two volumes per year

September, 1932
Volume XLI, No. 1

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND

JOURNAL OF
NETIC PSYCHOLOGY

Child Behavior, Animal Behavior,
and Comparative Psychology

EDITED BY
CARL MURCHISON

$7 00 per volume
Single numbers $4 00

QUARTERLY
Two volumes per year

September, 1932
Volume XLI, Number 1

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

SEPTEMBER, 1932

Chaque article est suivi d'un résumé en français
Jedem Artikel wird ein Referat auf deutsch folgen

THE CONFIGURAL FACTOR IN CHILDREN'S LEARNING . . . 3
By Harold E Jones and Dorothy Dunn

THE RELATIVE STRENGTH OF THE PRIMARY DRIVES IN THE WHITE RAT . . . 16
By C. J Warden

LEARNING AND MATURATION IN PRESCHOOL CHILDREN . . . 36
By Josephine Rohrs Hilgard

THE ASSOCIATION SPAN OF THE WHITE RAT . . . 57
By Lucien H. Warner

AN EXPERIMENTAL SEARCH FOR THE "CONDITIONED RESPONSE" . . . 91
By Lucien H Warner

SOME NEW BASES FOR INTERPRETATION OF THE IQ . . . 116
By Beth L Wellman

ISOLATED ACTION OF COMPOUND STIMULI IN A LOCOMOTOR HABIT OF RATS . . . 127
By Wayne Dennis and J M Porter, Jr

THE EFFECT OF MOTIVATION UPON INTELLIGENCE TEST SCORES . . . 136
By Julius B. Maller and Joseph Zubin

(over)

Worcester, Massachusetts

Entered as second-class matter August 3, 1897, at the post-office at Worcester, Mass., under Act of March 3, 1879

THE DETERMINATION OF A RELIABLE INTELLIGENCE QUOTIENT FOR THE YOUNG CHILD . . . 152
By Ruth Updegraff

"HABITS" OF THREE SUCCESSIVE NURSERY-SCHOOL GROUPS AND SOME RELATIONS BETWEEN TRAITS . . . 167
By Harriet E. O'Shea

GESTALT PRINCIPLES IN THE SIDEWALK DRAWINGS AND GAMES OF CHILDREN 192
By Lauretta Bender

SOCIAL FACILITATION OF FEEDING IN THE ALBINO RAT . . . 211
By H. F. Harlow

SHORT ARTICLES AND NOTES 222

Some Factors Influencing the Afternoon Sleep of Young Children (Ruth Staples) 222

A Critical Note on Some Recent Research on Mechanical Ability (John W. Cox) 228

APPARATUS 235

A Life-Size Alley Maze for Children (F. Davis and M. Batalla) . . . 235

BOOKS 240

Carl Murchison [Ed.], *A Handbook of Child Psychology* (Barbara S. Burks) 240

Books Received 219

THE CONFIGURAL FACTOR IN CHILDREN'S LEARNING*

From the Institute of Child Welfare of the University of California

HAROLD E JONES AND DOROTHY DUNN

INTRODUCTION

Considerable theoretical importance has been given to the familiar "relative choice" problem illustrated in the description of the following experiment, which was conducted by Kohler with a three-year-old child

> ". . . two boxes were placed before the child, one with a brighter and one with a darker color The child was told to take one and without further aid he soon learned to choose always the brighter box which contained candy, and to reject the other box which was empty. . . . When the child was able to make his choice virtually without error, the critical test was given [presenting the brighter together with a new and still brighter box] . . . the child invariably and without hesitation chose the new and brighter box" [Koffka (4, p 140)]

It is inferred that a relative choice would not occur if a child's phenomenal experience were built up of elementary sensations, or if the learning process consisted of a simple substitution of stimuli in specific S-R patterns. Since the behavior depends upon the "characteristics of a configuration," rather than upon the retention of "absolute positive qualities," Koffka believes that the phenomena involved in these experiments are configural It is easy to attach too much weight to these expressions. In the present instance, we need not assume that the term "configural" carries any special significance beyond that of "relative" or "relational."

If relative choice possesses the theoretical significance that has been claimed for it, a fuller quantitative study would seem to be desirable. The following questions arise,

1. At a given age level, what can be found regarding individual differences in this trait? It seems unlikely that relative choice is a uniform all-or-none reaction, inherent in the individual and independent of situational factors

*Received in the Editorial Office, April 20, 1932

2 What is the relationship between relative choice and chronological and mental age? In the terms of Koffka's discussion, configural reactions represent a genetically fundamental type of response; a tendency toward absolute choice would indicate a maturing capacity to analyze or isolate the factors in a situation, and hence should be correlated positively with other measures of maturity. On the other hand, both Kinnaman (3) and Johnson (2) appear to regard absolute responses as of a lower order than the relative type of reaction.

3. To what extent may relative choice be dependent upon extraneous variables in the procedure? In Kohler's experiment as described above, the darker and brighter boxes may be denoted as 2 and 3. The child was trained to choose Box 3. In the critical tests, he was presented with 3 and with 4, the latter being a new and still brighter box. A choice of Box 4, in the critical test, was interpreted as due to a tendency to relative choice. In the individual case, however, it may have indicated merely an attraction toward the specific new stimulus, either because of its novelty or because of its greater brightness. Or it may have indicated the rise of a specific inhibition against Box 3, which was able to take effect when the possibility of a new reaction was offered in place of the more strongly inhibited reaction to Box 2.

4. To what extent is relative choice dependent upon the test content? This involves the problems of specificity and of consistency in response. According to Koffka, relative reactions "become a compulsion" when the cues are qualitatively very similar to one another. Warden (5) has pointed out that any technique which favors the isolation of the specific stimulus also favors absolute reactions; this occurs, for example, when a standard stimulus is paired successively with various other stimuli.

Procedure

The present investigation was planned as an exploratory study rather than as a test of any single hypothesis. In order to provide a variety of test material, four sets of stimulus cards were used, each set presenting systematic changes in a single variable: the first, area; the second, form; the third, brightness; and the fourth, hue Each set was composed of four cards designated 1, 2, 3, and 4. Card 1 was respectively the smallest area, the thinnest rectangle, the lightest gray, and the yellowest orange. Card 4 was respectively the

TABLE 1

Card no	Area Area	Area Sides	Form Area	Form Sides	Brightness (grays)	Color
1	1 00"	1 00"	2"	8 x 2.5	Very light	Orange yellow
2	1 50"	1.25"	2"	9 x 2 2	Light	Yellow orange
3	2 25"	1 50"	2"	1.1 x 1.8	Medium	Orange
4	3 40"	1 85"	2"	1.2 x 1 67	Dark	Red orange

largest area, the most nearly square rectangle, the darkest gray, and the reddest orange (see Table 1).

After some preliminary experimenting with the size of areas in the first series, an inch square was selected for Card 1, with an increment of 50% in area for each succeeding card Likewise, after two inches was decided upon as an appropriate area for the form series, the proportions of the four figures were arranged from thinnest to broadest in such a way that the contrast between any two cards was clearly noticeable Standard Milton Bradley coated colored papers were used in the preparation of the figures in all but the brightness series For this series "Hering gray" papers were employed.[1] In order to have a more objective and reliable scale of grays than would be possible with one individual making the selection, a class in psychological measurements was asked to select from the entire series a very light, light, medium, and dark gray; the four papers finally chosen for the experiment were based on the average judgment of the class All the small figures in the four series were mounted on four-inch black cardboard backgrounds, the cards being faced with lantern-slide glass which could be wiped clean after handling.

Table 1 gives in compact form the characteristics of each of the 16 cards of the series.

The apparatus used consisted of a box 14 inches in length, with (1) a shelf for the presentation of the cue-stimuli, (2) two electric buttons, directly under the shelf, to be operated by the child after he has chosen one or the other of two stimuli, (3) a reward compartment, which automatically opens (exposing a reward) when the correct button is operated, (4) a multiple switch at the rear of the box, enabling the experimenter to establish connections between either

[1]The most probable positions for these papers in a series of 50 are 3, 10, 24, and 31 respectively See Brown (1)

button and the reward compartment, and (5), a housing for a dry-cell and the necessary wiring.

Cards 2 and 3 were used throughout the training for all subjects and all series. A training series consisted of six trials, followed immediately by a critical test. The length of the training series was determined in a preliminary experiment, the object being to employ a standard number of trials long enough to establish a complete discrimination habit in some but not all subjects. The correct or rewarded choice was, of course, varied in position in the six trials, the position schedule being right, left, left, right, left, right

On the basis of CA and MA, the sample was divided into four substantially equivalent groups (A, B, C, and D), working on the following schedule:

TABLE 2

Group	Cards presented in training series	Correct choice in training series	Cards presented in critical test	"Relative" choice
A	2 and 3	3	3 and 4	4
B	2 and 3	3	1 and 2	2
C	2 and 3	2	1 and 2	1
D	2 and 3	2	3 and 4	3

Group A was trained to select Card 3, the "larger than," "thinner than," "darker than," or "redder than" card of each pair throughout the entire series. The critical test was made with Cards 3 and 4; the latter, of course, to be chosen if a "relative" choice were made.

Group B also was trained on Card 3, but the critical test was made with Cards 1 and 2. Thus a child making the "relative" response in this series, i.e., choice of Card 2, would not be reacting to a new card but to what was actually the wrong response in the training series.

Groups C and D were trained to select Card 2, the smaller, the thinner and darker, and the yellower cards of the 2-3 pair. It may be noted that, while a "relative" choice may be made in the case of each group, the opportunity for a positive "absolute" choice is not offered in the case of Groups B and D. In some cases a non-relative choice might be viewed as an absolute avoiding reaction, through a persisting tendency to pick any stimulus *other* than the one which they had learned to avoid in the training series. The training, how-

ever, emphasized only the positive aspect: the reward for a correct reaction

The two cards for the first training series were already set up when the child was brought into the experimental room. After he was seated at a table with the problem-box facing him and the experimenter opposite him, the following explanation was given:

> "Now we are going to play a game. Watch what I do (here a Milton and Bradley colored lentil was dropped into the reward box). If I choose the *right* card and press the button *that belongs to it* (indicating the button just below the card), the box will open and I can get the prize. (Nothing happens) That must be the wrong card. If we had chosen the other card and pushed the button belonging to it, the box would have opened and you would have found the prize" (illustrating the correct choice).

A piece of heavy cardboard large enough to hide the apparatus from the view of the child was used between trials, whether or not an actual change of cards was made.

> "Now let's see if we can open the box with the first card we choose (pressing the button that was the wrong one previously). That's the right one this time. Sometimes this button opens the box, sometimes that button opens it You must watch these cards and find out which one will open the box"

As a means of trying out the procedure, a preliminary study was conducted with a group of nursery-school children, and the main experiment was then undertaken in the kindergarten classes of four schools.[2] These schools were selected primarily on a basis of social status, including inferior, average, and superior neighborhoods. Within each kindergarten the only selection employed was to pick children of white native-born parents This precaution was taken to eliminate the possibility of a foreign-language handicap in comprehending instructions, and to secure a group which could be more readily duplicated in later studies.

For the total group, the mean CA was 68 2 months, with an *S D* of 3 7. For 72 of these who had received the Stanford-Binet, the mean MA was 77 1 months, with an *S D* of 7.5 After the completion of the four series on 80 subjects, the entire experiment was repeated, 62 of the subjects being available for a second testing The

---

[2]Acknowledgments are due to Dr Virgil Dickson for his interest in making possible the collection of data in the Berkeley schools, and to the principals and kindergarten teachers of the John Muir, Lincoln, Washington, and Cragmont schools for their cooperation and assistance

interval between the two experiments was in every case at least eight days.

Results

In a total of 568 critical tests (both experiments, all series) a relative choice was made in approximately half of the instances ($54.3\% \pm 1.2$). While this suggests a chance determination, a closer inspection of the data will show the error of such an inference. On a basis of pure chance, the most probable frequency of relative choices would be 2, in a series of 4 critical tests. Only one child in 16 would be expected to make a maximum score of 4; a similar proportion would be expected to make the minimum score of 0. Table 3, based on 80 cases taking the first experiment and 62 cases taking the second, gives a distribution of relative (R) scores which is obviously determined by systematic factors opposing chance. More children than would be expected obtain a relative score of 4; also, more than would be expected obtain a relative score of 0, indicating a consistent non-relative or absolute choice.

The evidence is even more striking if we consider the percentage of children who make 8 relative choices in the 8 trials of the two experiments This is 9.6%, as compared with a chance expectation of .04% Nineteen and two-tenths per cent of the children make a relative choice either 7 or 8 times in 8 trials, as compared with a chance expectation of 3 5%. Fourteen and four-tenths per cent make a non-relative choice either 7 or 8 times in 8 trials, as compared with the same chance expectation of 3.5%. We cannot make the further inference, however, that the relative-reaction type is more common among these children than the absolute-reaction type; the apparent slight excess of relative over non-relative choices is due primarily to Groups B and D. In these groups, opportunity was lack-

TABLE 3
Percentage of Cases Making a Given *R* Score

R score	Percentage expected on a chance basis	Obtained percentages in 1st experiment $N=80$	Obtained percentages in 2nd experiment $N=62$
4	6.25	18.7	19 4
3	18.75	27 5	24 2
2	50 0	18.7	27.4
1	18.75	17 5	17 7
0	6.25	17 5	11.3

TABLE 4

PERCENTAGE OF CASES MAKING A "RELATIVE" CHOICE

	First experiment						Second experiment					
Group	N	Area	Form	Brightness	Color	Total	N	Area	Form	Brightness	Color	Total
A	20	55	55	55	60	56.2	11	36	45	54	64	50
B	20	60	60	45	50	53.7	18	67	61	39	61	57
C	20	50	55	35	55	48.8	17	65	53	41	47	51.5
D	20	60	55	50	50	53.7	16	56	81	50	56	60.7
Total	80	56.2	56.2	46.2	53.7	53.1	62	56	60	46	57	54.8

ing for a positive absolute reaction, since the non-relative choice was also a non-identical choice (see Table 2). For the two experiments taken together, the relative over the non-relative ratio was 56.44 for Groups B and D, 51.49 for Groups A and C.

Other than this not very significant difference, Table 4 reveals no definite effect of mode of procedure, or of repetition of the experiment. The effect of test content, however, is indicated by the results for the brightness series, in which there is a fairly consistently smaller proportion of relative choices. It will be shown later that this is due to a greater difficulty in establishing discrimination habits in this series.

If definite and persisting tendencies exist, different for different children, toward either relative or absolute types of response, it should be possible to demonstrate some degree of reliability in these responses. The correlation between R scores on the first and second experiment was $.35 \pm .075$. Applying the Spearman-Brown prophecy formula, the reliability of both tests taken together was found to be .52. While the obtained r for one test by the other is barely more than four times its $P.E.$, it is probable that this has suffered some attenuation through the small number of class intervals (maximum score 4) and the resultant effect of coarseness of grouping. In other words, a predicted reliability of .52 for the two experiments is probably smaller than that which would be empirically obtained, for in this case a test twice as long would also provide a more adequate score range. Another approach to the problem of consistency of response is given by Table 5, which indicates for 62 cases the per-

TABLE 5
PERCENTAGES MAKING CONSISTENT RESPONSES
($N=62$)

Series	Consistent relative choice in both experiments	Consistent absolute or non-relative choice in both experiments	Percentage expected on a chance basis
Area	40	29	25
Form	42	21	25
Brightness	24	34	25
Color	39	24	25
Average	36 2	27	25

centage of cases in which, for each series, a given type of response is shown in both experiments.

The indications are that the children who make a relative choice in the first experiment tend to be consistent to a degree greater than chance, and tend to be more consistent than those whose first choices are of the non-relative type; as previously pointed out, however, the latter effect is due to the B-D procedure, and is not revealed in equal degree among the children in Groups A and C.

The question now arises as to the determinants of these two types of response. Are they accompanied by differentiating characteristics in other traits? We may consider first the relationship to intelligence A priori, it is not clear that under the conditions of the experiment a positive correlation would be expected between MA and R score. While Koffka speaks of the identical or analytic factor in *reaction* as a product of higher development, and of the response to configurations as a very primitive type of mental achievement, on the other hand the tendency to look for and respond to relations is commonly regarded as an important constituent of intelligence, and the capacity for higher orders of configural organization is no doubt fundamental for intelligent behavior The actual correlations with CA and MA have proved to be near zero. For 62 cases, the total R score on two experiments correlated $18\pm.08$ with CA, and 29 ± 08 with MA. It remains to be determined whether more significant correlations would be obtained in groups covering a wider age range

A further correlation relevant to this problem is that between the R score and the "learning score" derived from the training periods of the present experiment A word of explanation is necessary with re

gard to this measure of learning. Since six trials were given in each of four series in each experiment, the learning score extended through a possible range of 0 to 24 (0 indicating no evidence of learning, 24 indicating no errors) A score higher than 20 would hardly be expected except through chance successes on the initial trials The actual score range was from 6 to 23, the mean being 12.6 on the first experiment and 14.7 on the second (a reliable gain). The self-correlation was .42, becoming .65 when stepped up by the Spearman-Brown formula to predict the reliability of the learning scores for the two experiments taken together The total learning scores show no significant correlation with MA (.15 ± .09). A substantial relationship, however, can be shown between learning score and R score, .47 ± .06 for the first experiment, and .56 ± .06 for the second. These coefficients are not reliably different from the self-correlations of the variables employed, while it is probably not legitimate in this case to correct for attenuation, the evidence is clear that a genuinely high community of function exists between learning to discriminate the cues in the training periods, and making a relative choice in the critical test.

It would appear that the more successful the training series, and the more easy the discrimination of stimuli, the greater the tendency to respond with reference to pattern rather than with reference to identical elements. As a specific example, if the child is shown areas 2 and 3, the more readily he learns to react to 3, the less likely he is to react to it when 3 and 4 are presented; and the more readily he has learned to avoid 2, the more likely he is to react to it when shown 1 and 2. This is also brought out in Table 6.

For learning scores of 4, 5, and 6 there is a definite trend in the direction of increased relative choice. There are several possible interpretations of this finding.

1 Children who have "just" learned, or incompletely learned, to make the necessary discrimination of cues, will give random reactions in the critical tests But by this explanation, the percentage of relative choices should not drop below 50, in the case of children whose learning scores are so poor as to suggest incomplete mastery of the task The fairly consistently low percentages found for children whose learning scores are 0, 1, or 2 indicate that they are not reacting at random in the critical tests, but are making choices which are systematically non-relative, in terms of our classification

2 A low learning score may not indicate a difficulty in estab-

lishing discrimination, on the contrary, it may accompany discrimination, with the response persistently directed toward the wrong cue. In such a case, a child trained on 2 and 3, and required to select 3 in order to obtain the reward, may repeatedly and systematically make the wrong selection In the critical test, if he is offered 3 and 4, and chooses 3, this would be listed as an absolute choice in our records, but would really indicate a relative choice on the part of the child. It is very difficult to know if this has actually occurred. Its importance is probably slight, as indicated by the fact that the incidence of low learning scores is approximately the same as that which would be expected by chance

3. The first theory implied that a low relative choice score is due to a low learning score. But causation may operate in the opposite direction A child who learns readily (in this experiment) may do so by paying attention to the total field. He reacts to differences, and to pattern. In the critical test he continues this type of reaction, and makes a relative choice The opposite type of response is shown by children who fixate on certain absolute characteristics of the stimulus, such as the general form of a square At first they react indifferently to either the small or the large square, since they are responding merely to "squareness." Later they isolate the additional factor of size. This results in slow learning and results also in an absolute choice in the critical tests.

4. A further hypothesis is that overlearning is favorable to configural reactions. In a series of a standard number of repetitions, the quick learner obtains more practice in the correct response than the slow learner Whether he has learned through the use of relative or of absolute cues, in the course of overlearning he responds increasingly to the relationships of the two stimuli

From the data now available, the writers feel unable to choose between the third and fourth of these interpretations. Additional experiments are necessary, in order to deal with each of these hypotheses specifically.

Conclusions

Referring again to the four queries presented at the beginning of this article, we may conclude.

1. Under the conditions of these experiments, wide individual differences exist among children in the tendency toward relative choice.

2 The tendency to relative choice is unrelated to mental or chronological age, within the range represented in this sample.

3. Relative choice persists with differing types of content, and with differing procedures in the critical tests.

4. The frequency of relative choice is related to the efficiency with which discrimination is established in the training series It is inferred that this indicates either (*a*) relative reaction as a characteristic of rapid learning, or (*b*) relative reaction as an accompaniment of overlearning.

In subsequent studies, it will probably be wise to increase the number of cases employed for any given procedure, to increase the number of types of subject-matter, and also to plan a schedule of variable overtraining in the period preliminary to the critical tests It may also be desirable to introduce variable delay periods before the critical tests.

The value of such studies, however, will be dubious if they are directed toward any such pretentious aim as the reconstruction of a psychological system More pertinent, and more promising for our understanding of children's learning, will be the formulation of problems dealing with the origin of tendencies to configural and absolute reaction, their degree of specificity, their dependence upon the situation, their modifiability under training, and their significance in complex learning situations.

References

1 Brown, W. Practice in grading and identifying shades of gray *Psychol Rev*, 1915, 22, 519-526

2 Johnson, H M. Visual pattern discrimination in the vertebrates *J. Anim. Behav*, 1914, 4, 319-339, 340-361, 1916, 6, 169-188

3 Kinnaman, A J Mental life of two Macacus rhesus monkeys in captivity *Amer. J Psychol*, 1902, 13, 94-148, 175-218

4 Koffka, K The growth of the mind New York, Harcourt, Brace, 1924 Pp xvi+383.

5 Warden, C. J, & Rowley, J. B The discrimination of absolute versus relative brightness in the ring dove, *Turtur risorius*. *J Comp. Psychol*, 1929, 9, 317-333

Institute of Child Welfare
University of California
Berkeley, California

LE FACTEUR DE FORME DANS L'APPRENTISSAGE DES ENFANTS

(Résumé)

On a préparé une expérience pour étudier (1) les différences individuelles dans le choix relatif et le choix absolu, (2) la relation entre le choix relatif et l'âge chronologique et l'âge mental, et (3) l'influence des variations dans les processus et dans le contenu des tests sur le choix relatif On a entraîné 80 élèves d'un jardin d'enfant à choisir une de deux cartes de grandeurs différentes, et ensuite on a fait subir des tests critiques, en présentant une de ces cartes avec une plus grande (ou plus petite) carte On a divisé les enfants en quatre groupes équivalents pour contrôler divers facteurs du processus On a fait subir des tests semblables avec d'autres types de matière (forme, clarté, couleur); le deuxième jour on a répété toute l'expérience On a trouvé que dans les conditions de ces expériences, il existe de grandes différences individuelles entre les enfants dans la fréquence du choix relatif Il persiste une tendance systématique au choix relatif (ou au choix non relatif) chez certains individus, malgré les différents types de contenu et les processus différents des tests critiques Parmi les enfants de l'âge du jardin d'enfants, la fréquence du choix relatif n'a aucune relation avec l'âge chronologique et l'âge mental Il existe cependant une grande relation entre le choix relatif et l'efficience de l'apprentissage dans la série entraînée. Deux interprétations alternatives sont que (a) le type relatif de réaction est un trait caractéristique de l'apprentissage rapide pour les matières spécifiques de test étudiées, (b) la tendance à la réaction relative devient plus grande pendant les périodes du nouvel apprentissage

JONES ET DUNN

DIE EINWIRKUNG DER GESTALT AUF DAS LERNEN BEI KINDERN

(Referat)

Es wurde ein Versuch eingerichtet zur Untersuchung (1) der individuellen Unterschiede in Bezug auf relative gegenuber absoluter Wahl, (2) des Zusammenhangs zwischen relativer Wahl und kronologischem und geistigem Alter, und (3) der Einwirkungen der Anderungen des Verfahrens und des Versuchsmaterials (test content) auf die relative Wahl (relative choice) Es wurden 80 vorschulpflichtige Kinder eingeubt, eine von zwei verschiedenen Grossen von Karten zu wahlen, und es wurden dann kritische Prufungen gemacht, in denen eine dieser Karten zusammen mit einer noch grosseren (oder einer noch kleineren) Karte angeboten wurde Die Kinder wurden in vier aquivalente Gruppen geteilt, um eine Kontrolle verschiedener Einwirkungen im Verfahren zu ermoglichen Es wurden mit anderen Sorten von Versuchsmaterial (in Bezug auf Form, Helligkeit, Farbe, u s w verschieden) ahnliche Versuche ausgefuhrt An einem zweiten Tag wurde die ganze Versuchsserie wiederholt Man fand, dass unter den in diesen Versuchen herrschenden Bedingungen, grosse individuelle Unterschiede unter den Kindern in Bezug auf die Haufigkeit der relativen Wahl bestehen Es beharrt bei gewissen Individuen eine systematische Neigung zur relativen (oder zur nicht-relativen) Wahl, ob die Arten des Versuchsmaterials oder die Verfahren in den

kritischen Prufungen verschieden sind oder nicht Innerhalb des Kindergartenalters steht die Häufigkeit der relativen Wahl zum kronologischen oder zum geistigen (mental) Alter nicht in Beziehung Die relative Wahl und die Tuchtigkeit des Lernens in der Übungsserie stehen aber zu einander in enger Beziehung Zwei alternative Deutungen der Befunde sind: (*a*) dass die relative Reagierungsart eine Eigenschaft des raschen Lernens darstellt, in Bezug auf die Versuchsgegenstande die spezifisch untersucht wurden, (*b*) dass die Neigung zur relativen Reagierungsweise wahrend der Perioden des Überlernens verstarkt wird

JONES UND DUNN

THE RELATIVE STRENGTH OF THE PRIMARY DRIVES IN THE WHITE RAT*[1]

From the Animal Laboratory, Department of Psychology, Columbia University

C J. Warden

Introduction

The experiment to be discussed this evening has been designated in the announcement as "a biological study." In order that this qualification lead to no misunderstanding, perhaps I should say that the present study is neither more nor less biological than the typical experiment in comparative psychology today All of our work aims to be broadly biological rather than narrowly psychological, and this trend is distinctly on the increase.

It is true that this has not always been the case. Not so long ago the comparative psychologist was interested primarily in the subjective life of infra-human organisms. He studied the behavior of the animal with the explicit intention of interpreting its mental states or of determining its mental level. If his observations and experiments did not lend themselves to mentalistic analysis, he was likely to feel that he was wasting his time. In fact, it was once common to offer an apology in connection with the publication of behavior studies when the experimenter did not feel warranted in hazarding mentalistic interpretations along with the facts presented. But happily this point of view has largely passed away. During the last two or three decades the comparative psychologist has become much more modest than formerly. He no longer harries himself with the inscrutable mystery regarding the nature of the animal mind He has become convinced that, even if an animal possesses a mind—whatever a mind may be—we can never gain any genuine insight into the subjective experience of infra-human types. They cannot introspect —after the human fashion—and hence cannot tell us how they feel, and perhaps we humans would not understand them if they did.

*Received in the Editorial Office, November 23, 1931.

[1]Read before the Biology Section of the New York Academy of Sciences, American Museum of Natural History, November 9, 1931

The comparative psychologist of today is primarily interested in the behavior rather than the mind of the animal. He seeks to observe and measure behavior and to explain complex activities in terms of simpler activities He has ceased to confuse the issue by attempting to infer anything whatsoever as to the possible conscious concomitants of behavior His aim at present is to become a good naturalist instead of remaining a bad psychologist. He attempts to take up the task of biological analysis where the morphologist and physiologist leave off. In general, the morphologist deals with structure, the physiologist with the functions of organs and groups of organs, while the comparative psychologist studies the behavior of the organism as a whole in its larger environment To put it otherwise, physiology is primarily concerned with relatively simple behavior and comparative psychology with the more complex behavior of the organism. If the problems of the newer animal psychology are determined to some extent by those of human psychology, it may be said that its approach to these problems is as objective as that of physiology. Please believe, therefore, that I am not harboring a lot of psychological interpretations in my own mind in attempting to keep the discussion strictly biological. I am dealing with the problem precisely as I might before the Psychology Section of the Academy—I am keeping nothing back.

Among the numerous lines of analysis in comparative psychology none would seem to be of greater genuine importance than that which concerns itself with the dynamics of behavior. It is clearly necessary to know something of the springs of action in a given type if we are to understand even its most simple activities. The basic explanation of behavior must be found, in the last analysis, in the internal mechanisms which are more or less characteristic of the species Although these mechanisms may be set off by external stimuli, the nature of the response is largely determined—in complex types at least—by the mechanisms themselves. This internal organization is evolutionary and hereditary in origin, and serves to explain why a dog reacts one way to a certain stimulus and a cat another way to the same stimulus. When a given internal system is aroused by an appropriate stimulus, the resulting behavior is characteristic of the species within the limits of individual variation It is usual to regard types of behavior which seem to be very definitely tied up to one or another of these internal systems as instincts. Moreover, the instincts are generally looked upon as comprising the dynamic factors underlying all of the behavior

of the organism—the springs of action that serve to make the living individual a going concern

While there are few who would wish to challenge the fundamental implications of this point of view, there is a growing resentment against the use of the term instinct among comparative psychologists. It would be entirely out of place for me to enter into this field of controversy this evening I mention the matter merely because I wish to make use of the term which most comparative psychologists now employ instead of instinct—i.e, the term *drive* We speak of the sex drive, the hunger drive, the thirst drive, and the like, rather than of the sex instinct, hunger instinct, thirst instinct, etc. The term *drive* is preferred, among other reasons, because an instinct is usually regarded as manifesting a more or less specific pattern of behavior. Anyone who has observed the natural activities of a complex organism, however, knows that much of the behavior usually classed as instinctive is not definitely patterned. When a cat is hungry, for example, it is likely to make use of numberless behavior patterns in order to secure food. Many of these same patterns may be called out under sex excitation and the like. It thus appears that the so-called instincts cannot be differentiated from one another in terms of specific behavior patterns Natural movements, such as climbing, clawing, running, jumping, etc., if these lie within the repertory of the individual, may occur in connection with any of the "instinctive" activities of the type

Now the term *drive* preserves the essential element of the concept of instinct without implying the occurrence of specific and invariable behavior patterns in each case It emphasizes the importance of the several internal mechanisms of behavior as evolutionary and hereditary factors definitely related to biological needs These represent the primary dynamic factors in all behavior They constitute the more stable components of activity, while the repertory of responses called out in any case often involves a considerable amount of variation and overlapping. The primary drives should be thought of as the innate and persistent cores of behavior which may become attached from time to time to any of the sensory-motor systems of the organism However, I do not wish to press the point. I shall use the term *drive* in the present discussion and those of you who prefer the older concept of instinct may substitute the latter without doing any particular violence to the facts to be presented

As might be expected, the analysis of the drive-complex of a species

or variety offers much greater difficulties than the usual laboratory experiment in the animal field In the first place, there is the seemingly insurmountable problem of securing measures of the strength of the several drives in terms of comparable units How can hunger, thirst, sex, and the like, which are so intrinsically different, be reduced to a common scale of objective behavior? What test of strength of drive that is equally fair to each can be applied under like conditions to all? The whole matter is greatly complicated by the fact that the several drives must be interdependent, since the underlying mechanisms are known to be interrelated For example, the degree of thirst manifested after a given period of water deprivation is in part a function of the kind and amount of food supplied to the animal during the time In like manner, the strength of the sex drive is influenced by the conditions utilized to induce hunger and thirst Finally, it should be borne in mind that the animal under test is likely to be highly excitable because of the fact that the internal dynamic mechanisms are under definite stimulation and examination. This suggests the need of every precaution to prevent outside distractions and emotional disturbances during the testing process itself. Perhaps the best way to indicate the difficulties involved, as well as one manner of resolving them, will be to enter at once into the discussion of the method and procedure employed in the present study Aside from more than a year spent in the development of the method, the project here reported covered a period of some four years and included the testing of more than 1000 white rats The project[2] was supported financially to the extent of $5000 by the Council for Research in the Social Sciences of Columbia University The white rat was used on account of its small size and availability, and because enough is known of the white rat in laboratory situations to enable us to standardize conditions rather definitely for this type

The Method

The apparatus and procedure finally adopted for use in the project was devised by Jenkins and Warden and is known as the Columbia Obstruction Method A general notion of the lay-out of the apparatus can be had by referring to the ground plan as shown in Figure 1. It consisted essentially of a long box divided into three com-

[2]For a complete report on this project, see *Animal Motivation Experimental Studies on the Albino Rat*, by C J Warden and collaborators (1)

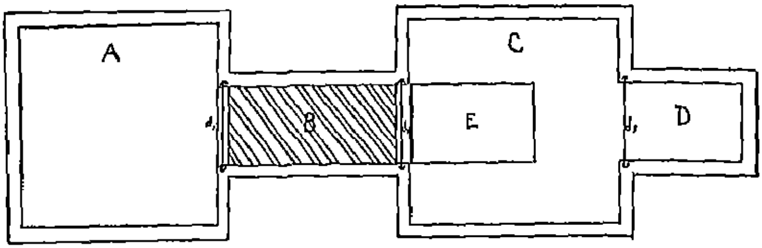


FIGURE 1

DIAGRAM OF FLOOR PLAN OF THE OBSTRUCTION BOX

A, entrance compartment, *B*, obstruction compartment; *C*, *D*, divided incentive compartment; *E*, release plate, *d*, automatic door (operated by release plate) between two divisions of incentive compartment

partments. The first and last of these were approximately 10 inches in all dimensions while the middle compartment was a narrow, low tunnel, about 10 inches long, extending over an electric grid. The test animal was placed in the first compartment, an appropriate incentive in the last compartment, while the electric grid in the middle one constituted the obstruction or obstacle over which the animal must pass in order to reach the incentive beyond.

The use of an obstacle in testing the intelligence of animals is, of course, common enough in the animal laboratory. The principle had been applied under somewhat different conditions by Morgan, Moss, and others to the measurement of drives. But in previous work the shock was varied with incentive-drive conditions in order to find out how much shock would balance off a given degree of hunger, thirst, etc. In the Columbia obstruction apparatus, however, the shock was kept constant for all drives and thus served as a common measure of the strength of each of them. The general principle here is precisely that of the human intelligence test—individuals and conditions become directly comparable in terms of scores made in performing an identical task. The only factor which was varied, throughout the entire project, was the series of drive-incentive conditions to be studied. Differences in strength of drive were indicated by differences in the number of partial and complete crossings of the grid, under the various conditions, during a test period of 20 minutes.

It seems necessary to pass over a number of minor but exceedingly important points of technique very briefly. Each detail of the procedure finally adopted was determined by preliminary experimenta-

tion, rather than by arbitrary decision or convenience, and the data relating to the development of the method have appeared in published reports A 20-minute test period was utilized because it was found that a longer period gave no better index of strength of drive. As a matter of fact, our results show the same ranking of drives when only the scores of the first half (10 minutes) of the 20-minute period are considered The animals were all of the same strain—Wistar Institute stock direct, or first generation reared from these in our own laboratory. This is an inbred strain and hence more homogeneous in hereditary background than the animals carried by commercial dealers. The sexes were reared together up to 150 days of age so as not to disturb the normal development of the sex drive All groups were segregated at 150 days to eliminate pregnancies, allowed 5 weeks to become accustomed to our specific laboratory conditions, and tested at the average age of 185 days, the range of variation in age being 3 weeks (175 to 196 days). At this age, the white rat is fully mature and in top condition for tests of this sort. The standard shock, used throughout the project, was determined empirically for the particular weight of animal to be tested The grid was supplied with alternating current of 60 cycles, with terminal pressure of 475 volts, external resistance of 10,100,000 ohms, and current of 0 047 milliamperes by means of an elaborate electro-stimulator. In addition to the fact that a high voltage current is generally advisable in stimulation work on animals, it has the specific advantage of eliminating possible differences in stimulation value arising from individual variations in skin resistance In spite of these large figures —or perhaps because of them—the shock was very slight, being just enough to make the rat lift his feet a bit gingerly in crossing the grid

A few points relative to the test procedure should be noted. The animals had become used to handling and, in general, were very tame before being tested Immediately preceding the 20-minute test period, each animal was allowed to cross the uncharged grid to the incentive beyond four times in succession in order to arouse the appropriate drive. On the fifth crossing, the grid was charged thus bringing the obstacle into play The 20-minute test period began immediately thereafter, and crossings and partial crossings were recorded minute by minute. Since the partial crossings add nothing to the analysis of the results, we shall simplify matters by referring only to complete crossings in the later discussion. The experimenter was eliminated from the test situation by a device based upon the principle of the one-

way light screen Tests were made only at night—between the hours of 9 P M and 4 A M.—during which time the white rat exhibits its maximum activity McCollum's standard diet for the white rat, plus a regular supply of greens, guaranteed against any vitamin deficiency which might have disturbed certain of the drives. The incentives placed in the last compartment were samples of those used to satisfy the several drives under the conditions of normal cage life The satisfying of the drive by partaking of the incentive after each crossing was limited as follows (1) hunger—nibble of powdered food, (2) thirst—licking moist nipple of water bottle, (3) preparatory sex behavior but without copulation, (4) maternal—attention to litter, and (5) exploratory—brief random activity A period of only 30 seconds was allowed in the incentive chamber. The purpose of the incentive was to keep the drive aroused and to heighten it by continued stimulation It is clear that, if the animals had been allowed free and unrestrained access to the incentive after each crossing, satiation and a waning of the drive would have taken place rather soon.

The central aim of the project was to carry the analysis far enough to make possible a ranking of the five primary drives investigated. hunger, thirst, sex, maternal, and exploratory. Such a comparison could not be made unless the maximum strength of each drive should be determined. To find the maximum strength of each drive it was necessary to test a wide range of internal conditions in each case, so planned as to include the optimum condition. Since the optimum status could not be known in advance, the schedule of conditions to be tested in each case had to be arranged as the work proceeded. The range of internal conditions actually tested in establishing the maximum strength of each of the several drives was as follows: (1) hunger—0, 2, 3, 4, 6, 8 days of starvation, (2) thirst—0, 1, 2, 4, 6 days of water deprivation, (3) male sex—0, 6, 12 hours, and 1, 4, 7, and 28 days of sex-deprivation; female sex—oestrum (cornified stage) and 7 stages of dioestrum; (4) maternal—first litter (young) and multiparous (standard age), and (5) exploratory—one group only tested This range of conditions did not yield, necessarily, the optimum condition for the maternal drive, but since this drive stood highest as tested the matter of ranking the five drives is not disturbed thereby The exploratory drive scored so low in the one test, given under very favorable conditions, that it did not seem likely that it would rank other than lowest had the tests been continued indefinitely.

In order to standardize sex conditions in the males, each period of sex deprivation was begun after the sex drive had been thoroughly satiated by repeated copulation with a female in heat during a two-hour stretch. The oestrus condition of the female was determined by a histological examination of a vaginal smear taken after the animal had been tested, the individual being then thrown into the group indicated by the smear. Permanent slides of the smears were made and our classification was checked by Professor Papanicolaou of the Cornell Medical School. In no case was an animal tested more than once, and the standard group for each drive condition investigated consisted of 20 animals. Only a single animal, of course, was tested at any one time.

As mentioned above, the several drives are more or less interdependent by reason of the interrelations which exist among the internal mechanisms underlying them. This makes the task of isolating one drive at a time for testing a very difficult matter. In the strict sense, isolation of such factors is probably not possible in an organism as complex as the white rat. The best that can be done is to insure the *dominance* of the drive to be tested by keeping all of the others as quiescent as possible at the time. A summary of the controls used are shown in Table I, and a more detailed account will be found in Part VII of the volume mentioned in the footnote above (1). In general this was accomplished by devising a set of standard conditions for quiescence in each of the drives and always applying four of these as indicated when testing the fifth or aroused drive. The principle may be illustrated by stating the conditions maintained with respect to the other four drives while the sex drive was being tested: (1) hunger was allayed by allowing the animal access to food in the living cage up to the moment of test, (2) thirst was allayed by supplying the living cages with water up to the time of test, (3) the maternal drive was eliminated by using non-pregnant females, and (4) the exploratory drive was kept quiescent by allowing the animals to be normally active in the cages up to the time of test. The same general scheme, with necessary shift in the specific internal factors to be kept quiescent, was applied to each of the other drives as well. By arousing in the animal only one biological need at a time, and placing before it the incentive related to this need, the dominance of a given drive for the moment was insured. This was, of course, necessary if we were to secure the maximum indices in each case. How well we succeeded in this attempt to keep all drives except the one under test quiescent

TABLE 1

SUMMARY OF EXPERIMENTAL AND TEST CONDITIONS FOR NORMAL DRIVES

Drive condition tested	Size of group M.	F	Comb	Incentive (Compartment D)	Incentive response	Method of keeping other drives quiescent: Hunger	Thirst	Sex	Maternal	Exploratory
Hunger										
0 days	10	10	20	Sample of regular diet (McCollum's mixture) in compart.	Nibble of powdered food, 30 sec. limit	Drive being tested	Regular water supply to time of test	Males seg 35 days, Females in dioestrum	Only non-pregnant females without litters	Specific incentive objects present, no place to explore
2 days	10	10	20							
3 days	10	10	20							
4 days	10	10	20							
6 days	10	10	20							
8 days	10	10	20							
Thirst										
0 days	10	10	20	Regular water bottle in compart.	Moistening tongue on damp nipple, 30 sec. limit	Regular diet except greens omitted	Drive being tested	Same as above	Same as above	Same as above
1 day	10	10	20							
2 days	10	10	20							
4 days	10	10	20							
6 days	10	10	20							
Sex—male										
0 hours	20			Female in cornified stage, except compart empty for control group	Nipping, biting, nosing genitalia, mounting, etc, 30 sec limit	Regular diet to hour of test	Same as thirst above	Drive being tested	Same as above	Same as above
6 hours	20									
12 days	20									
1 day	20									
1 day (control)	20									
4 days	20									
7 days	20									
28 days	20									

TABLE 1 (*continued*)

Drive condition tested	Size of group M	F	Comb	Incentive (Compartment D)	Incentive response	Method of keeping other drives quiescent: Hunger	Thirst	Sex	Maternal	Exploratory
Sex—female										
Early congestive		7		Normal male, except compart. empty for control group	Nipping, biting, nosing genitalia, exposing vulva, 30 sec. limit	Same as above	Same as above	Drive being tested	Same as above	Same as above
Oestrum (cornified)		21								
Same (control)		22								
Late cornified		3								
Post-ovulative		6								
Dioestrum (4 stages)		32								
Maternal										
First litter		9		Litter in maternity cage	Attending to litter, 30 sec. limit	Same as above	Same as above	Females with litter	Drive being tested	Same as above
Multiparous (stand age)		10								
Exploratory										
Males (stand age)	20			Exploration box	Random activity, 30 sec. limit	Same as above	Same as above	Males seg 35 days	Males only	Drive being tested

is a question not so easily answered However, the controls employed to effect this were uniform throughout the project and hence one drive could not have been favored over another. The relative isolation of each drive at the time of testing was thus approximately the same so that the indices obtained may be regarded as typical and comparable throughout

Ranking of Normal Drives

The central purpose of the project, as previously stated, was to analyze the normal drive-complex of the white rat in a manner that would make it possible to rank the several drives tested in order of strength Such a ranking must be made, of course, on the basis of the maximum indices obtained for each drive The latter are given, together with the usual statistical values, in Table 2 The maximum scores would seem to represent the highest dynamic value of the several drives under optimum conditions of motivation. The scores for the two sexes show no significant differences in so far as the hunger, thirst, and sex drives are concerned, and hence may be com-

TABLE 2
SHOWING MAXIMUM SCORES (CROSSINGS) FOR THE VARIOUS DRIVES, ARRANGED IN RANK ORDER

Drives tested	Condition of maximum drive	Size of group M	F.	Comb	Average	Standard deviation	Coefficient of variation
Male							
Thirst	2nd day	10			21 10	11 60	54
Hunger	4th day	10			19 10	5 87	31
Sex	1st day	20			13 45	4 03	30
Exploratory	Only 1 tested	20			6 00	4.89	81
Female							
Maternal	Standard age		10		22 40	9 14	41
Thirst	1st day		10		19 70	11 10	56
Hunger	3rd day		10		19 00	8 91	47
Sex	Oestrum		21		14 14	5 14	36
Combined							
Maternal	Standard age		10		22 40	9.14	41
Thirst	1st day	10	10	20	20.40	11 40	56
Hunger	3rd day	10	10	20	18 20	7 58	42
Sex	1st day & oestrum	20	21	41	13 80	4 64	34
Exploratory	Only 1 tested	20			6 00	4 89	81

TABLE 3

SHOWING THE RELIABILITY OF THE DIFFERENCES BETWEEN THE MAXIMUM SCORES (CROSSINGS) OF THE DIFFERENT DRIVES WITH SEXES COMBINED

Drive conditions tested	S.D. of difference	Difference between averages	Difference / S.D. of difference	Chances in 100 of a true difference
Hunger and thirst	*3.06*	*2.20*	*0.72*	*76*
Hunger and sex	1.84	4.40+	2.39	99
Hunger and maternal	3.35	4.20	1.25	89
Hunger and exploratory	2.02	12.20+	6.03	100
Thirst and sex	2.65	6.60+	2.49	99.4
Thirst and maternal	3.85	2.00	0.52	70
Thirst and exploratory	2.77	14.40+	5.19	100
Sex and maternal	2.98	8.60	2.89	99.8
Sex and exploratory	1.31	7.80+	5.95	100
Maternal and exploratory	3.09	16.40+	5.31	100

Legend. The plus sign indicates that the first drive of the pair has the higher score.

bined. The maximum indices for the several drives when these scores have been combined are shown in Table 2 and on the graph of Figure 2.

As may be seen, the maternal drive clearly takes first rank, showing an average of 22.40 crossings during the 20-minute test period. Further evidence of the strength of the maternal drive is furnished by the much higher score (28.33 crossings) of the younger, or first-litter group. On the other hand, the exploratory drive, the incentive for which was a large playground attached to the apparatus, ranked lowest of all, averaging only 6 crossings during the test period. It is clear, therefore, that the opportunity to explore does not constitute a very strong incentive. This incentive seemed to arouse a genuine drive, however, since the score is higher than that made by several control groups which crossed to the third compartment of the apparatus when no special incentive was present. From a general survey of the maximum scores as indicated in the graph of Figure 2 (also Table 2), it appears that the drives may be ranked in the following order: maternal, thirst, hunger, sex, and exploratory.

However, the matter is not quite as simple as it seems. As may be noted, the amount of difference between the maximum scores is rather small in some cases. The question arises as to whether or not the obtained differences are true and dependable in every instance. In

answering this question we must take into account the range of individual variation within the groups as well as the amount of difference between the two group averages. When the difference between group scores is small, and the variation among individual scores within the various groups is large, it often turns out that the apparent difference between the group scores is not reliable. We must examine our results with respect to this point, therefore, before drawing any final conclusions. The computations covering this matter of the reliability of the differences between the maximum scores for the several drives are shown in Table 3 According to the usual rule, a difference between two averages is not certainly reliable unless such a difference is approximately three times the standard deviation. In applying this rule, we will make only those comparisons that relate directly to the proposed series of rankings, since these are the ones in which we are especially interested at the moment. It will be simpler if we state the question of reliability in terms of the chances in 100 that the differences found are true differences. We find that the chances are 70 in 100 that maternal is higher than thirst, 76 in 100 that thirst is higher than hunger, 99 in 100 that hunger is higher than sex, and 100 in 100 that sex is higher than exploratory. This means that when we compare the maximum scores for a given drive with the one next below it, the differences are not large enough to be statistically reliable except in the case of the last two comparisons, i e., hunger *vs.* sex drive and sex *vs.* exploratory drive It is reasonable to suppose, however, that if larger groups in each case had been tested the measures of reliability might have been somewhat higher than they are.

Without calling in question the generally accepted criterion of reliability which we have just applied, we may make use of another which seems to be more to the point. This second method of evaluating the reliability of differences between averages is similar to the first except that it deals merely with ordinal position. It is known as the method of combining averages and is recommended by the English statistician, Yule. As a matter of fact we are concerned, here, not so much with intercomparisons among the group averages, as with the question of the ordinal position of each average in the series as a whole. We want to know whether or not we can say that the order indicated on the graph of Figure 2 is the true order of this series of averages. The computations involved in applying this criterion of reliability are given in the lower section (combined) of

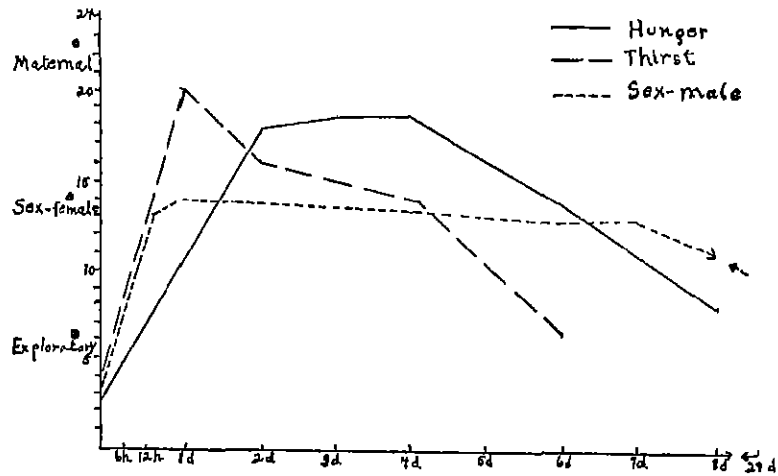


FIGURE 2

CURVES REPRESENTING THE DIFFERENT DRIVE-CONDITIONS TESTED, WITH THIRST AND HUNGER SCORES FOR MALE AND FEMALE COMBINED

The several periods of deprivation are indicated on the abscissa and the average number of crossings on the ordinate

Table 4 When this method is applied we find that the chances are 100 in 100, or perfect, within limits of a small fraction which may be safely ignored (lowest 99 6), that the ranking of the maximum scores indicated on the graph is a true and reliable ranking This means that, in so far as we have secured sound indices of the several drives, we are warranted in ranking them as to maximum strength as follows. maternal, thirst, hunger, sex, and exploratory The question of the soundness of our analysis of the drive-complex of the white rat depends ultimately, of course, upon the soundness of our method In so far as our general procedure and controls were adequate to this type of measurement, our results may be regarded as conclusive. Perhaps I may be permitted to add that we are now engaged in extending the method with suitable modifications to the rhesus monkey

STUDIES OF SPECIAL CONDITIONS

In addition to the analysis of the normal drive-complex of the white rat just reported, a study was made of a number of special conditions in connection with the project The same general method of testing

TABLE 4

SHOWING THE RELIABILITY OF THE DIFFERENCES BETWEEN MAXIMUM SCORES, AS RELATED TO RANK ORDER

The difference between a given score and that for each drive of lower rank, taken singly, is treated in columns 2, 3, 4, and 5, in the last column the scores for all the drives of lower rank have been combined

Drive tested	Rank order	Chances in 100 of a true difference greater than 0 Thirst	Hunger	Sex	Exploratory	Combination
Male						
Thirst	1		69	98	100	99 4
Hunger	2			99 7	100	100
Sex	3				100	100
Exploratory	4					
Female						
Maternal	1	72	80	99 6		96
Thirst	2		56	89		86
Hunger	3			95		95
Sex	4					
Combined						
Maternal	1	70	89	99 8	100	99 6
Thirst	2		76	99.4	100	99 7
Hunger	3			99	100	100
Sex	4				100	100
Exploratory	5					

was used in each of these investigations. However, the results secured cannot be directly compared with those previously discussed since certain special conditions were introduced in each case for the purpose of disturbing the normal drive in some manner. Nevertheless, these studies are pertinent to the topic of the evening inasmuch as they throw light upon the nature of drives in the white rat It will be unnecessary to do more than state very briefly the special conditions introduced in each instance and mention the main result. I shall limit myself to the following three studies which seem to me to be of most interest. (1) the effect on the strength of the hunger drive of delaying access to the incentive or food, (2) the effect on the male and female sex drive of segregating the sexes for various periods of time, and (3) the effect on the male and female sex drive of gonadectomy, vasotomy, and injections of placental and orchic extracts. These will be discussed separately and in the order indicated

In all of the tests previously described, the animal obtained momentary access to the incentive as soon as it had crossed the grid each

time. We were now interested to see whether a drive would be weakened if the animal were restrained from access to the incentive for a short period of time after each crossing. Hunger was selected as the particular drive to be tested in this respect, and a 48-hour starvation period was used since this had been shown to induce the maximum strength in this drive. The manner of restraining the animal was simple enough. The food was placed as usual in the little box marked *D* on the diagram of Figure 1, but the door separating this box from compartment *C* was left closed during the period of delay The animal merely remained in the latter compartment—often in a state of great activity—until the door was raised. The following series of delay intervals was used. 0 seconds, 15 seconds, 30 seconds, 1 minute, 3 minutes. A group of 20 animals was tested for each delay interval, making 100 in all.

It was found that a delay as short as 15 seconds decreased the *number of crossings during the 20-minute test period approximately 43%.* Longer delays of 30 seconds and 1 minute did not further decrease the drive index, but the 3-minute delay brought about a decrease of 73% from the score of the zero, or control group. It should be stated, moreover, that the time comprising the delay intervals was counted out so that the delay groups were given the full opportunity of 20 minutes to cross We cannot explain just why an interval of only 15 seconds diminishes the strength of the hunger drive so much, since the animals remained quite active until the door was raised. It was noted, however, that the animals of the 3-minute group tended to become inactive after a time, and it seems possible that an interval of this length may have disturbed the normal association between the act of crossing and the food At any rate it is clear that the hunger drive remains at maximum strength only when immediate access to the incentive is to be had

The main purpose in the study of the segregation of the sexes was to determine whether or not such conditions would disturb the normal development of the sex drive We could not arrange for complete segregation in the strict sense of eliminating all sensory stimulation between the sexes The males and females were merely kept in different cages and so placed as to exclude visual stimulation as far as possible Our conditions prevented tactual stimulation, and, of course, sexual behavior of any sort, but certainly not olfactory influences Groups of males and females were segregated at 30 days, 150 days, and 185 days of age—all of them being tested at the standard

age (185 days). Copulatory activities in the white rat usually make their appearance at about 50 days, the age range, as at present known, being from 37 to 72 days. The three groups tested, therefore, represented pre-pubertal, post-pubertal, and adult age levels. Separate groups of males for each age level were tested to the female in oestrum, to the female in dioestrum, and to another male. In like manner, separate groups of females for each age level were tested to an active male, to a female in oestrum, and to a female in dioestrum. The use of three types of incentive in each case was necessary to determine the specific kind of disturbance which had been induced by the segregation, if any such had occurred. The male test animals were in splendid health, and the female test animals were in the cornified stage of the oestrus cycle as determined by vaginal smears. In general, it was found that the sex drive was fairly normal in strength in the post-pubertal and adult groups. However, both males and females that had been segregated before puberty (30 days) exhibited definite homosexual tendencies. That is, the males would cross more often to males than to females, whether in oestrum or dioestrum, while the females would cross more often to females, either in oestrum or dioestrum, than to active males Moreover, the animals segregated at this age level had developed much of the usual repertory of sex play in their cage life and manifested such homosexual activities in the apparatus after crossing the grid. When animals segregated at 30 days were allowed to mate freely with an active member of the opposite sex for some hours before the test, the homosexual tendency was less strongly exhibited. In general, the results seem to indicate clearly that, in the absence of the opposite sex from the pre-pubertal period onward, the sex drive may become definitely attached to a stimulus of the same sex. This does not occur, however, if the animal has been allowed to mate with the opposite sex during the developmental period Incidentally, the sex life of the male white rat varies greatly from individual to individual. We found certain virile males who, while very exciting to most females, were inclined to be highly selective in mating with females. Other males who seemed less vigorous and less stimulating were relatively promiscuous as to sex play, although mating was rarely consummated

The facts secured as a result of segregation suggested a further study of the sex drive by the usual operative techniques, such as gonadectomy, vasotomy, and the like. Males were castrated and females spayed at 150 days of age and tested in the drive apparatus

at the standard age (185 days) The scores were compared with those of suitable control groups which had undergone a sham operation at the same age. As might have been expected, the sex drive in both male and female gonadectomized animals was very greatly reduced. This reduction was even more apparent in a group of males castrated at 90 days of age. The drive was practically absent in males castrated at 30 days of age, this group showing, on the average, approximately 6 crossings. A double injection of placental extract (cow) prepared for the follicular hormone, given 72 and 48 hours before the test, served to restore the drive in both sexes very nearly to normal. An injection of castrated males with orchic extract was without effect upon the depleted drive Vasotomy performed upon 14-month-old males which had become somewhat senile—crossing only about 7 times during the test period—was ineffective in restoring the sex drive. In general, the results of these correlations between definite physiological conditions and behavior tend to show the intimate relation between drive activity and specific internal mechanisms

Perhaps the chief value of the project as a whole lies not so much in the facts discovered regarding the white rat, as in the perfecting of a method which now can be applied to most of the higher animals The importance of the analysis of the primary dynamic factors in behavior has long been recognized, but no method was at hand by which the intercomparison of drives could be carried out. For this reason, this general field has been neglected by the comparative psychologist in so far as systematic experimentation is concerned. Let us hope that this neglect will now be remedied by a vigorous program of research, the primary aim of which should be to determine the underlying motivation of living organisms along strictly objective lines

Reference

1 Warden, C J, Jenkins, T N, Warner, L H., Jenkins, M, Hamilton, E L., & Nissen, H W Animal motivation experimental studies on the albino rat New York Columbia Univ Press, 1931 Pp. 502.

Columbia University
New York City

LA FORCE RELATIVE DES IMPULSIONS PRINCIPALES CHEZ LE RAT BLANC

(Résumé)

Dans ce rapport il s'agit de l'analyse du complexe des impulsions chez le rat blanc au moyen de la méthode d'obstruction de Columbia, dérivant d'une étude d'une durée de quatre ans récemment complétée au laboratoire de l'Université de Columbia Les cinq impulsions testées ont l'ordre de rang suivant duquel on montre la constance (1) l'impulsion maternelle, (2) la soif, (3) la faim, (4) le sexe, et (5) l'impulsion exploratrice. On a basé cet ordre sur les résultats maxima obtenus pour une impulsion donnée dans les conditions optima de stimulation On n'a stimulé qu'une seule impulsion à la fois et on a maintenu tranquilles les quatre autres impulsions pendant qu'on a testé celle-ci. On a obtenu ainsi les résultats maxima dans des conditions qui ont rendu certaine la dominance de l'impulsion testée dans la situation Ceci le rend possible possible de placer les impulsions d'une façon definie dans leur position d'ordre dans le complexe général des impulsions chez le rat. On a mesuré la force de l'impulsion par le nombre de fois que l'animal a traversé un grille électrisée au stimulant de l'autre côté Le choc et toutes les autres conditions sauf le stimulant ont éte maintenus constants pour toutes les impulsions. Outre l'analyse principale, on a étudié de nombreuses conditions spéciales telles que la ségrégation, la gonadectomie, des injections de l'extrait placentaire, un stimulant retardé, etc L'auteur cite un livre par lui-même dans lequel il a mis les résultats détaillés

WARDEN

DIE RELATIVE STÄRKE DER ELEMENTARTRIEBE BEI DER WEISSEN RATTE

(Referat)

Der Verfasser erstattet hier Bericht uber eine Analyse des Triebkomplexes (drive-complex) der weissen Ratte, durch Gebrauch der Columbia Hemmungsmethode (Columbia Obstruction Method), die sich aus einem vierjahrigen, neulich im psychologischen Laboratorium der Universitat Columbia vollendeten Arbeitsplan entwickelt hat. Die funf geprüften Triebe nehmen die folgende, als zuverlassig erwiesene Rangordnung an: (1) Muttertrieb (maternal), (2) Durst, (3) Hunger, (4) Geschlechtstrieb, (5) Erforschungstrieb Die Rangornung grundete sich auf die hochsten, durch einen Trieb unter moglichst gunstigen Umstanden der Hervorrufung erzielten Zahlen Es wurde jedes Mal nur ein Trieb hervorgerufen, die ubrigen vier Triebe wurden wahrend der Prufung dieses einen im Ruhestand gehalten Es wurden auf diese Weise die hochsten möglichen Zahlen erhalten, unter Umstanden die versicherten, dass der zu prufende Trieb in der Situation vorherrschend war. Es wird auf diese weise die definitive Rangordnung der Triebe in Bezug ihre Ordnungsstellung (ordinal position) im allgemeinen Triebkomplex der Ratte ermoglicht. Die Starke eines Triebes wurde gemessen an der Häufigkeit mit der das Tier ein elektrisches Gitterwerk kreuzte, um daruber hinaus den Reiz zu erlangen Die Starke des Schocks und alle ubrigen Umstände mit Ausnahme des Reizes wurden durch alle Reize hindurch konstant erhalten

Es wurde nicht nur die grundsatzliche Analyse gemacht, sondern man untersuchte auch die Einwirkungen zahlreicher spezieller Bedingungen, wie z B Isolierung, Gonadectomie, Einspritzungen mit Mutterkuchenextrakt (placental extract), Verzogerung der Reizung, u s w. Es wird auf ein Buch vom Verfasser hingewiesen, worin die einzelnen Befunde umstandlich zusammengefasst worden sind. (Warden, C J. Animal Motivation. New York Columbia Univ Press, 1931 Pp 502)

WARDEN

LEARNING AND MATURATION IN PRESCHOOL CHILDREN*

From the Psychological Laboratories of Yale University

JOSEPHINE ROHRS HILGARD

The present study,[1] for which a group of Merrill-Palmer nursery-school children were subjects, was undertaken in the hope that it might yield further information concerning the relative importance of maturation and practice in the development of motor skill in young children. For this purpose, it appeared that abilities which were just beginning to develop in the children offered the best opportunity for investigation. Since previous studies had shown that the abilities of buttoning, cutting with scissors, and climbing a ladder appear in children between the ages of 24 and 36 months, these three skills were selected for study. A control group and a practice group were used in order to test the effects of maturation and general practice (control group) as against the effects of intensive special training (practice group) during a three-month period.

RELATED STUDIES

It is generally recognized that the improvement of an ability with age is due both to the maturation of innate growth factors and to the cumulative effects of functional experience. To discover the rôle which each plays, and to estimate the degree of interdependency, is a problem which has been approached primarily from the field of animal behavior, and more recently from that of child behavior.

The classical experiments of Spalding (15) and of Shepard and Breed (14) on maturation in the flying of birds and the pecking of

*Recommended by Edward S Robinson, accepted for publication by Carl Murchison of the Editorial Board, and received in the Editorial Office, November 12, 1931.

[1]The author is indebted to Dr Rachel Stutsman of the staff of the Merrill-Palmer School for direction in the organization of this study and in the preparation of the report, which is based on an essay submitted to the faculty of the Graduate School of Yale University in partial fulfillment of the requirements for the degree of Master of Arts in psychology, 1930. Acknowledgment is also due Professor Edward S. Robinson of Yale University for numerous suggestions.

chicks are familiar. More recently, Carmichael (3) has shown that drugged embryos of frogs and salamanders, remaining motionless during a period of growth, when denarcotized soon swam so well that there was difficulty in distinguishing them from the group which had been swimming five days. It appears that the rapid learning of the delayed group was possible because they were able to take advantage of a greater maturity.

In his careful investigations over a period of years, Coghill (4) correlated the behavior of Amblystoma with the structural development of the nervous system. His results clearly indicated that behavior, which represents a progressive expansion of an integrated pattern, develops in a sequence of movements consistent with the order of development of the nervous system. Coghill concludes that the *form* of the behavior pattern is determined by laws of growth within the organism, the normal experience of the animal with reference to the outside world appears to have nothing specifically to do with it. On the other hand, in determining when, and to what extent, the potentiality of behavior shall be expressed, there is an interaction between the processes of growth and the experience of the individual. Thus specificity of function is fixed by the relations into which the innate elements grow, and by the excitation from the environment Whether the conclusions drawn from the Amblystoma apply to the higher vertebrates and man is an open question. Coghill believes that it is reasonable to suppose that in their broad outlines they do.

The problem of the relative influence of maturational growth forces and learning on human development has been approached from a number of standpoints. Such familiar instruments as the Developmental Schedules of the Yale Psycho-Clinic (7), the Merrill-Palmer Scale of Mental Tests (17), and the Terman Revision of the Binet-Simon (18) have been organized on the assumption that we can differentiate levels of development in the individual from birth to maturity That such development proceeds independently of environment, except where the latter deviates too much from the normal, is also inferred. Of environmental conditions which might be expected to influence development to some extent, we distinguish, for the sake of clarity, those which offer opportunities for the general exercise of developing abilities from those which offer opportunities for the specific practice of a specific ability

Studies of the effect on development of opportunities for general

exercise have been made by Blackhurst (2), Hildreth (13), Goodenough (9), and Barrett and Koch (1) With the exception of Blackhurst, whose study was confined to the value of play apparatus for motor control, the object of these experiments was to measure the effect of nursery-school training on the intelligence test scores of young children. To what extent would the early optimal opportunities for acquaintance with many materials affect general development? The conclusions, while in general tending to minimize the influence of this training on mental test scores, are somewhat at variance Hildreth found that children with nursery-school training, upon entering the first grade, showed only a temporary superiority, and Goodenough's results tended to confirm this. In her experiment the advantage of the nursery-school group over the control group on the second examination was slight enough to be within the limits of chance. On the other hand, Barrett and Koch, using a nursery-school group and a control group of orphan children, found that after nursery-school training the IQ's of the nursery-school group had risen from 91.71 to 112.57, while those of the control had risen from 92.59 to 97.71, or approximately half as much An analysis of the nature of the nursery-school activities leads them to conclude that direct practice effects do not account for this rise.

Like the attempts to measure the effect of general exercise on development, a number of studies directed toward the problem of determining the effect of specific practice of a specific function have utilized certain developmental items for which standardized norms are available. This procedure seems desirable because these norms provide a fairly accurate idea of the status of the ability in question. At the same time, the experimenter may be criticizing the significance of these norms, for if practice shows much improvement in the ability tested by certain items they may prove unsuitable for inclusion in a developmental scale, in spite of age differentiation.

To test the hypothesis that intensive practice preceding full maturity may stimulate and increase the rate of growth of certain capacities, Gates and Taylor devised two tests, one on memory for oral digits (5), which has been standardized as part of the Stanford-Binet Intelligence Test, and one on speed of tapping (6) They matched a practice group of children in chronological age (four to six years), mental age, intelligence quotient, and ability in the initial performances of the activity to be tested. They found that after

78 days of practice in memory for oral digits the practice group had gained 2.07 digits, the amount gained by the average untrained child in six years, according to the Stanford-Binet Test The control group had gained 0.67, or one-third as much. Particularly interesting is the fact that a retest four and a half months later showed that the advantage of the practice group had been entirely lost and that the two groups were as nearly equal as at the beginning of the study. Similar results were obtained in the tapping experiment. Gates and Taylor conclude that maturational processes, continuing in both control and practice groups, are not perceptibly affected by special training.

Gesell and Thompson (8) point out the possibilities of studying maturation offered by the method of co-twin control They report a study of monozygotic twins 46 weeks of age For six weeks Twin T was given daily practice in climbing and cube behavior, while Twin C, the control, who had no training during that period, was given daily practice the following two weeks. At 55 weeks of age, the climbing ability of the twins was nearly the same; that is, Twin C had acomplished as much in two weeks of practice as Twin T had accomplished in six weeks. From these data Gesell and Thompson infer that the superior performance of C, with only a third the opportunity for practice, must be due to maturation of the processes involved. Further, it was impossible to demonstrate any significant influence of training upon the cube behavior patterns of Twin T Strayer's study (16) on the vocabulary development of the same pair of twins was interpreted to mean that, in the field of language, a maturational difference of even five weeks had a definite influence on the relative effectiveness of training

Other studies on the efficiency of learning at different levels have dealt particularly with the development of specific skills. Goodenough and Brian (10) tested the development of skill in throwing rings over a post in three groups of preschool children Over a period of 50 days, Group A practiced with no instruction, Group B received some instruction, and Group C received thorough instruction. It was found that during this period Group A (10 children) had progressed only 11 5 in comparison with 17 5 for Group B (6 children), while Group C (4 children) had made the tremendous gain of 42 5 It would be interesting to have the results of a fourth control group having neither practice nor instruction, and also to have the results of a retest a few months later

In a similar investigation, Hicks (11, 12) found that maturation and undirected practice in young children aged two and one-half to six and one-half years were more important in the development of the complex motor skill of hitting a moving target than was systematic, well-motivated practice once a week for eight weeks. In a parallel study the same children were given strength, perforation, and tracing path tests once during the period when the initial target tests were given and again three months later. The results showed that increase in skill on these tests without specific practice was comparable to increase in skill on the target test either with or without specific practice. The author believes, therefore, that improvement in skill may result from factors other than specific practice, such as the influence of structural maturation and of general practice.

That improvement in skill does result from factors other than specific practice there can be little doubt. The problem is one of *how much* the environmental factors can contribute. It seems now that more experimentation on abilities which are in the process of development will add to our knowledge of the total picture.

The Problem

The present experiment, which utilized a group of Merrill-Palmer nursery-school children as subjects, was attempted in the hope that it might throw further light on the problem of maturation and learning in young children. The skills of buttoning, cutting with scissors, and climbing a ladder, which previous observations had shown to be developing in children between the ages of 24 and 36 months, were chosen for study.

Clues in regard to buttoning and cutting performances at this level appear in the Merrill-Palmer Scale of Mental Tests (17). According to this scale, the age at which the average child is first able to button a one-button strip is 30.5 months, the age at which he can first button a two-button strip is 33 months; and the ability to cut gashes is an average performance for children from 24 to 29 months of age. Though few children of this age are able to make a series of successive cuts in paper, there is so striking an improvement in this ability between the ages of 30 and 35 months that 32% of 50 children of this age tested in the standardization of the scale were able to cut such strips. In contrast to buttoning and cutting,

which involve the development of power and skill in fine motor coordinations, climbing brings into play gross motor coordination

Experimental Procedures and Materials

Two groups of 15 children each were matched according to chronological age, mental age, sex, and approximate initial ability in the three skills—buttoning, cutting with scissors, and climbing. Because of sickness, only 10 children in each group completed the experiment. Table 1 gives the means for each group of children.

Although the first tests were made chiefly in December, a few of the children were tested in November, and a few the first week in January. The mean initial test for both groups falls in December Dating from the first week in January, the practice group received intensive training for 12 weeks Table 2 shows the number of practice periods each child had.

After the first three weeks of training, the group was retested, and thereafter retests were made at two-week intervals when possible. At the end of the twelve-week period the control group was retested and then received intensive training for four days. One child in the control group developed measles after the second day of training, and one was unable to return for more than one day Eight chil-

TABLE 1

Group	Mean CA	Mean MA	Sex
Practice	28 3	29 1	6 girls, 4 boys
Control	28 6	29 9	6 girls, 4 boys

TABLE 2
Practice of Groups, Number of Practice Periods, Including Retests

Child	Buttoning	Cutting	Climbing
Kathryn	26	29	42
Mary E	24	26	37
Marilyn J	27	30	46
Jean	24	29	41
Cynthia	26	30	44
Frances	28	29	45
John	24	27	—
Patrick	25	25	38
Daniel	24	28	37
Stanley	27	29	36

dren, therefore, completed the entire intensive training program, while the other two completed enough of the program to permit the inclusion of their data During this week of intensive training for the control group, the practice group received no specific training, but were tested at the beginning and again at the end of the training period.

In the following discussion of the techniques employed in testing the three skills, a description of the testing and retesting periods is in each case followed by a statement concerning the practice period.

Climbing

1 *Testing the skill attained in climbing.* This first test consisted of climbing a three-step ladder 2½ feet high, stepping on to a table at the top, and then climbing down again. On the table were various toys which might attract and please the child These toys were changed frequently so that they did not lose novelty, and were varied to please individual children Some of the boys, for instance, responded particularly well to engines, while other children liked toy animals. The usual formula was, "Let's see how quickly you can climb up on the table today," and "Now let's see how quickly you can climb down."

Going up the ladder, the child was timed from the second he put one foot on the first step until he had both knees or feet on top. Going down, he was timed from the second his knees were in the middle of the top of the ladder until both feet were on the floor again. The child's final score was obtained by averaging the time required for climbing up and that required for climbing down.

The pleasure the children experienced in the climbing test brought them into a cooperative frame of mind for the cutting and buttoning. They liked coming into the room to see what we had that was "new," and they particularly liked being on top of the table. After the child had examined the toy there, he was told he could hand it to the experimenter and she would put it on the little table (where he would next cut and button) while he climbed down. When this method was followed, the child would sit down at the little table ready to do the cutting and buttoning. Very often he was allowed to play with the toy for a short time before the next material was introduced That the children retained their interest in the three tests is probably due very largely to this technique.

Only eight of the ten pairs of subjects completed this experiment.

TABLE 3

Group	Mean CA	Mean MA	Sex
Practice	28 3	29 1	5 girls, 3 boys
Control	28 6	29.9	5 girls, 3 boys

One child in the practice group was fitted with leg braces to straighten his knock-knees, and one in the control group did not take the initial test The two groups of eight each were matched in chronological age, mental age, and sex, as shown in Table 3.

In plotting the learning curve for the practice group, it was necessary to interpolate to determine the score of the children at definite two-week intervals for, though no child had had less than six retests, these did not fall at the same time for all the children. For the control group, the curve was drawn without interpolation.

During the practice periods, the experimenter tried in every way to help the children eliminate wrong methods and improve the time and quality of their performance, for example, in training the child to climb faster, the experimenter showed him how to put one foot, instead of two, upon each step of the ladder In other respects the practice periods were similar to the test period, except that the performance was not timed.

Buttoning

1 *Testing the skill attained in buttoning.* Usually the buttoning test was given immediately after the climbing test In this test the child's performance was scored according to the degree of his success or the length of time he took to button strips offering different degrees of difficulty The strip easiest to button was a folded piece of cloth with four buttons and corresponding buttonholes and a fifth button at the top already buttoned, so that the four buttonholes were in place over the buttons Two of these buttons were ¾ of an inch in diameter and two ⅞ of an inch If a child buttoned one button, he was given a score of one point; if he buttoned all four, he was given a score of two points Performances on the other button strips have been standardized as a part of the *Merrill-Palmer Scale of Mental Tests* The material consists of three pairs of 3x6-inch flannelette strips, the first pair having one button and buttonhole, the second pair two buttons and buttonholes, and the third pair four All these buttons are ⅝ of an inch in

diameter That the child can button the four-button folded strip described above before he can manage the pair of strips having only one button is explained by the element of difficulty added when the two strips are separate and the buttonholes must be placed over the button before the child can button the strips.

The method of presenting the material to the child is practically identical with that used in the Merrill-Palmer Scale, and the reader is referred there for a full account (17). The folded four-button piece was presented first, and then, in turn, the one-button, two-button, and four-button strips. No assistance was ever given in a test situation. The scoring shown in Table 4 closely follows that standardized in the Merrill-Palmer Scale.

The higher the number of points earned, the better is the performance rated.

Nine of the ten pairs of subjects were included in this test. The tenth pair was omitted from the results for several reasons. Though Bernice, the tenth child of the control group, matched Cynthia of the practice group in cutting and climbing, she scored 8 of the possible 13 points in the initial buttoning test, while Cynthia had a score of zero points. Further, when Bernice was retested in April, she scored 6 points on the test and could not score more than 7 points at any time. Since one atypical case, like that of Bernice, can noticeably affect the results when the number of cases is so small, and

TABLE 4
BUTTONING SCORING DEVICE

Material	Accomplishment	Points
Folded 4-button piece	1 button	1
	All buttons	1
1-button strips	Success	1
2-button strips	Time: 170" or less	1
	Time. 50" or less	1
	Time: 34" or less	1
	Time: 30" or less	1
	Time 23" or less	1
	Time 19" or less	1
4-button strips	Any success	1
	Time: 76" or less	1
	Time: 51" or less	1
	Time 42" or less	1
	Total	13

TABLE 5

Group	Mean CA	Mean MA	Sex
Practice	30.5	31.6	5 girls, 4 boys
Control	29.9	31.6	5 girls, 4 boys

since the two children were never well matched in this skill, this pair was omitted.

The two groups of nine children each were matched as follows (Table 5):

The mean curve for the practice group was obtained by interpolation, as described in connection with the climbing test

2. *Practice period.* During the practice period, the children had a wider variety of buttoning material from which to choose. A blue piece with two sizes of colored buttons down the sides made a very nice bag when all the buttons were buttoned. A red piece with black and white buttons in the form of a cross proved attractive to the children. There was also a blue folded piece with small colored buttons During the practice period the child could select the ones he liked, though he was urged in various ways to make his choice from actual test materials. The experimenter assisted as much as was necessary and at the same time acceptable to the child. In practicing with the strips, for example, the younger children had to be coached on the method of getting the buttonhole exactly over the button

Some of the children enjoyed buttoning; for others it was necessary to make the buttoning part of a game. One child who felt no fondness for buttons liked to put the buttoned strips into the baggage car of a toy train (secured in the climbing experiment) and send the train off on imaginary trips Another child pretended they were blankets for the animals and after buttoning the strips put them on the animals to keep them warm Though the interest in buttons was far from spontaneous toward the end, a fair degree of interest was maintained in ways such as these, varying much with the individual child.

Cutting

1. *Testing the skill attained in cutting.* The cutting test was given after the buttoning test To secure a measure of the child's performance and progress in this test and to provide a basis for scor-

ing, graph paper was used for the test material. A sheet of 8x10½ graph paper was cut vertically into two equal parts. On one of these were drawn two vertical red lines, 10 centimeters high and parallel to each other, on the other a line rising to a height of 14 centimeters was drawn at an angle of 45° to the base of the sheet. The child was given the sheet with the two vertical red lines and told to cut as carefully as possible along one of them, the experimenter indicating the point where he was to begin

This cutting test held considerable interest for the children. Sometimes the parallel vertical lines on the graph paper were supposed, when connected, to form a door After the child had cut along one of the lines, the experimenter would cut across the top to the other side and fold the piece back so that it looked like an open door and supported the paper At other times, the cut paper was called a house, the experimenter cutting a "smokestack" from the top after the child had finished cutting. Other similar devices were used to hold the child's interest.

Since there is no test sufficiently similar to this to prove of assistance in scoring the child's performance, it was necessary to devise a scoring system On the basis of the children's performances, a scale of successive levels was devised, ranging from 100 points for the poorest performance—no cutting at all—to zero for perfect adherence to the red-line length. Deviations from the red line were computed by counting the number of two-millimeter squares between the red line and the cutting. This total was then divided by the height attained A system of weighting was soon found to be necessary, since it was obvious that the chances of any degree of deviation from the line were considerably less when the child could cut only 4 centimeters along the line than they were when he could cut the whole length of the line. The method of scoring is shown in Table 6.

The data for all the ten pairs of children were available in the cutting test. The learning curve for the practice group was obtained by interpolation, as in the other two tests.

2. *Practice period.* During the practice period the graph sheets just described and picture pages torn from magazines proved interesting as practice material. If the child was about 24 months of age, he had to be taught how to hold the scissors and coordinate the movements of the two hands, if a little older, how to make sev-

TABLE 6

CUTTING. SCORING DEVICE

No manipulation	100 points
Opens scissors with both hands	95
Opens scissors with one hand	90
Cuts if paper is held for him	85
Cuts gashes Holds paper himself	75
Cuts 3-3.9 centimeters	$\frac{^{*}D \quad 250}{^{*}L}$
Cuts 4-4.9 centimeters	$\frac{D \quad 200}{L}$
Cuts 5-5.9 centimeters	$\frac{D \quad 150}{L}$
Cuts 6-6.9 centimeters	$\frac{D \quad 100}{L}$
Cuts 7-7.9 centimeters	$\frac{D \quad 50}{L}$
Cuts 8-8.9 centimeters	$\frac{D \quad 20}{L}$

*D = Deviation, L = Length

eral successive gashes in the paper, and finally, he had to be taught how to guide the scissors along the lines.

RESULTS

At the end of the 13 weeks covered by the study, when the practice group had been trained for 12 weeks and the control group had been given a final week of intensive training, the outstanding result is a marked similarity in the gains made by the two groups. Though the practice group leads in the gain made in cutting and buttoning, the groups are practically the same in climbing. In the buttoning test the practice group gained 21.4 points and the control group 15.6, or 73% as much. In cutting, the practice group gained 47.5 points and the control group 40.4, or 85% as much. The difference between the two groups in climbing is slight and unreliable—9.2 for the practice and 9.6 for the control.

It is evident that the two groups were very evenly matched in all three initial skills, for only in cutting is there a difference of even slight reliability. In April, after the 12 weeks of intensive practice, there is a reliable difference between the two groups in all skills, and at the time of the final retest the practice group has a lead of low reliability over the control in cutting and buttoning, but the difference in climbing is negligible.

TABLE 7

SUMMARY PROGRESS OF THE PRACTICE AND CONTROL GROUPS IN CLIMBING, CUTTING, AND BUTTONING

	Initial test December	After 12 weeks April	After 13 weeks April	Total gains
Buttoning				
(N = 9 in practice and control groups.)				
Practice group Mean ± *P.E.M.*	6.6±2 1	30 5±3 9	28.0±3 7	21 4±3 4
Control group Mean ± *P.E.M.*	7 7±2 1	19 0±2 6	23 3±3.1	15 6±2 2
Difference between means	1 1	11 5	4 7	5 8
P.E. of diff.	2 9	4.7	4 8	4.1
Diff./*P.E.* of diff.	0.38	2 47	0 98	1 41
Reliability of difference	Unreliable difference	Fairly reliable difference	Unreliable difference	Low reliability
Cutting				
(N = 10 in practice and control groups.)				
Practice group Mean ± *P.E.M.*	64 3±4 5	19 3±5 6	16.8±5 0	47.5±3 8
Control group Mean ± *P.E.M.*	73 6±4 4	40.3±8	33.2±7 2	40.4±4.8
Difference between means	9 3	21.0	16 4	7 1
P.E. of diff.	6 3	9 8	8 8	6 1
Diff./*P.E.* of diff.	1 48	2.14	1 86	1.25
Reliability of difference	Low reliability	Fairly reliable difference	Low reliability	Low reliability
Climbing				
(N = 8 in practice and control groups.)				
Practice group Mean ± *P.E.M.*	17 8±1 08	8.8±9 5	8 6±.73	9.2±1.33
Control group Mean ± *P.E.M.*	18 4±1 76	12.9±.70	8 8±.75	9 6±1 11
Difference between means	0.6	4.1	0.2	0.4
P.E. of diff.	2.06	1 18	1.05	1 73
Diff./*P.E.* of diff.	0.29	3 47	0.19	0.23
Reliability of difference	Unreliable difference	Fairly reliable difference	Unreliable difference	Unreliable difference

Certainly, the remarkable relative gains of the control group, with so limited an amount of training, suggest that factors other than specific training contributed to the development of these three skills. This is borne out by the fact that between the initial test and the first retest in April, before the week of specific training, the control group had gained 57 3% of its total score in climbing, 72.5% of its total in buttoning, and 82 4% of its total in cutting A comparison of the total gain made by the control group with the gain made by this group during the one week of intensive training also throws some light on the nature of the factors involved We find that in climbing 42 7% of the total gain comes during this week, in buttoning, 27.5%, and in cutting, 17 6%. These gains are all considerably in advance of those made by the practice group during any one week of its 12-weeks' training period In fact, the practice group in climbing required all of the previous month to make 42.7% of its total points, a feat accomplished by the control group in one week. Thus, it would appear certain that the rate of learning was accelerated toward the close of the 13-week period covered by the investigation.

The learning curves (Figures 1-3) for climbing are particularly significant in offering clues for the interpretation of the gains. An examination of the curves reveals positively accelerated learning for the practice group up to the last retest in March, learning being much more rapid toward the end of the period On the supposition that the control group lags behind the practice group throughout (a fair assumption because it is behind on every test), it follows that the curve of improvement for the control group is also positively accelerated between the initial trial and the first retest, for without positive acceleration it would cut across the practice curve, as in the figure This positive acceleration shared in by both practice and control groups we may interpret as signifving more favorable maturaton and general developmental potentiality for climbing toward the end of the period The practice group has taken a pronounced lead by the second week in April, showing that training effectively capitalizes this potentiality, and the very rapid improvement of the control group during the following week is further evidence in the same direction. That the two groups at the end are almost identical in their mean scores, the one after 12 weeks of practice, the other after one week of practice, shows the importance of the underlying factors These curves therefore demonstrate more

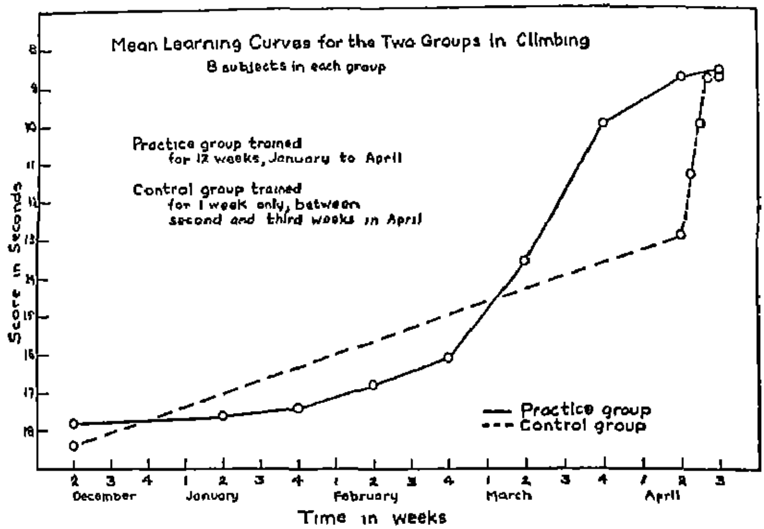


FIGURE 1

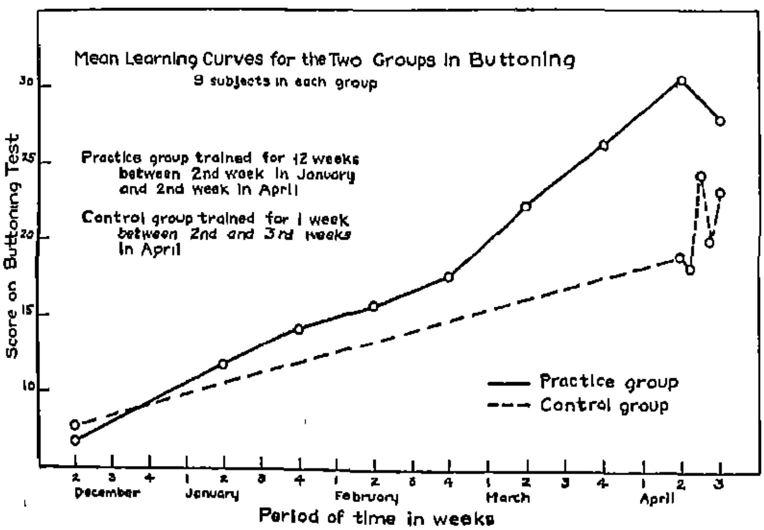


FIGURE 2

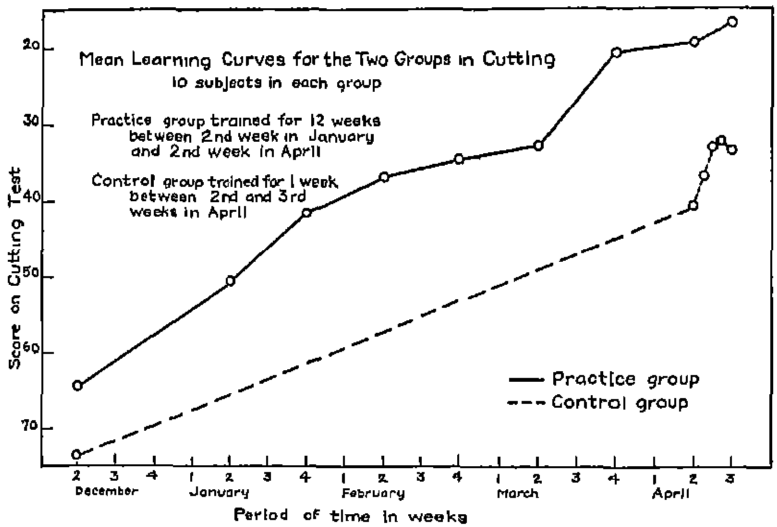


FIGURE 3

clearly than the others both the importance of the underlying factors and the temporary acceleration which can be introduced by special training. The decrease in acceleration of the practice group after the end of March suggests that there may be a limit to improvement during any given developmental stage, which is, of course, to be expected.

Where the curve for climbing shows only a slight gain during the first few weeks of training, the learning curve for buttoning rises steadily from the beginning, with a marked acceleration at about the middle of the training period continuing until the end of the period During the 13th week, when the practice group had no training, this curve shows a decline of several points Whether this drop is to be attributed to the discontinuance of training or to the fact that the children's interest in buttons had reached the saturation point is not clear. In view of the findings of Taylor and Gates, one may conclude that the gains made in a developing skill like buttoning as the result of special training are unstable in character and that the ability tends to resume its former level after a period of disuse However, since the data for cutting and climbing do not show a similar drop after the end of the practice period, it may be that the

length of the period of disuse necessary to bring about this drop in the level of ability achieved may vary both with the skill itself and with the degree of skill attained.

A consideration of the data on cutting and buttoning yields some information concerning the values of these two items as part of a mental test battery. As was mentioned earlier, even though certain skills show age differentiation, they may not be desirable in such a battery if, at the same time, they are much affected by special training. If it had been possible to retest the practice and control groups at a later period the results bearing on this point would be more clear cut. Nevertheless, it is evident from the slight differences between the two groups at the end of the experiment that special training, though it had some effect, was far outweighed by the general developmental factors. The fact that the curve for buttoning had already started to decline lends credence to the view that the slight superiority attained in this skill may have been temporary.

In interpreting the findings of this experiment it must be remembered that the control group was given a retest and four training periods during the one week of intensive training, while the practice group averaged only two training periods a week during 12 weeks, consequently, the frequency of the practice periods may be one factor in determining the increase in the rate of learning evident at the close of the study. It is impossible, also, to eliminate the factor of specific practice in interpreting the gain made by the control group during the period of no specific training. Yet children from 23 to 34 months of age are unlikely to be required to button their own clothes, and they are too young to take much interest in cutting unless they have considerable supervision and assistance. Climbing is usually a favorite activity, but whether a child does much of it at home depends on a number of conditions, such as the kind of play space available and the attitude of the parents toward allowing the child to climb.

In all skills, however, even though we can say that it is unlikely that the children have had much practice in these specific skills in the home, we cannot rule out, as Carmichael did by drugging his frog and salamander embryos, the factor of constant practice in related manipulatory activities. The child is continually handling objects and gaining greater skill in coordinating the fingers and the two hands as well as gaining skill in gross motor movements. Consequently, when he is retested after four months, it is to be expected

that he will have made progress quite aside from the specific training he has been given. Whether or not we shall term this progress "maturation" depends largely on our definition of the word Gesell and Thompson concluded, from the results of their study of twins mentioned earlier in this paper, that maturation was the dominant factor in the climbing ability of the twins, but, of course, the twin observed as a control, like the children of the control group in the present study, was gaining much general practice in activities related to his learning to climb at a later time—kicking, balancing, and the like

In this experiment, we cannot certainly distinguish between the gain to be attributed to maturation alone and that due to maturation plus practice in activities related to the specific skills studied What does appear is that maturation, plus this related general practice, accounts for the great gain made between the initial test and the initial retest of the control group, and that specific training throughout the twelve-week period was a far less important contributing factor in the development of these three abilities than was this general developmental trend.

Summary

Two groups of ten children each, aged 24 to 36 months, in the Merrill-Palmer Nursery School were equated for chronological age, mental age, sex, and approximate initial abilities in three skills: buttoning, cutting with scissors, and climbing The practice group was given 12 weeks of practice, with retests at two-week intervals. The control group, without special practice in the meantime, was tested at the end of the 12-week period, and then given four days of intensive practice. The practice group and the control group were both tested again after this period The initial differences between control group and practice group were unreliable on all the tests. After the 12 weeks of practice, the practice group exceeded the performance of the control group on all the tests, but one week of practice by the control group was sufficient to bring the scores of the control group and the practice group to similar levels The rapid relative gains of the control group are interpreted to mean that factors other than specific training contributed to the development of these three skills, factors which may be partly accounted for by maturation and partly by general practice in related skills. There is evidence from the learning curves of the practice group in buttoning and climbing

that improvement is more rapid in the latter part of the training period, consistent with the accelerated learning of the control group at the end of the experiment.

REFERENCES

1. BARRETT, H E., & KOCH, H L The effect of nursery-school training upon the mental-test performance of a group of orphanage children *J Genet. Psychol*, 1930, **37**, 102-122.
2. BLACKHURST, D J The value of play apparatus for developing motor control in preschool children. Iowa City State Univ. Iowa Library, 1927. Also abstract *28th Yrbk. Nat Soc Stud Educ*, 1929, 472-473
3. CARMICHAEL, L. The development of behavior in vertebrates experimentally removed from the influence of external stimulation *Psychol. Rev.*, 1926, **33**, 51-58
4. COGHILL, G E Anatomy and the problem of behavior New York Macmillan, 1929. Pp. xii+113
5. GATES, A I, & TAYLOR, G A An experimental study of the nature of improvement resulting from practice in a mental function *J. Educ Psychol*, 1925, **16**, 583-592
6. ———. An experimental study of the nature of improvement resulting from practice in a motor function *J. Educ Psychol*, 1926, **27**, 226-236
7. GESELL, A The mental growth of the preschool child. New York: Macmillan, 1925 Pp 447
8. GESELL, A, & THOMPSON, H. Learning and growth in identical infant twins An experimental study by the method of co-twin control. *Genet. Psychol Monog.*, 1929, **6**, 1-124.
9. GOODENOUGH, F L. A preliminary report on the effect of nursery school training upon the intelligence test scores of young children. *27th Yrbk Nat Soc Stud. Educ*, 1928, Pt. 1, 361-369.
10. GOODENOUGH, F L, & BRIAN, C R. Certain factors underlying the acquisition of motor skill by children of preschool age *J Exper Psychol*, 1929, **12**, 127-155
11. HICKS, J A. The acquisition of motor skill in young children *Child Develop*, 1930, **1**, 90-105
12. ——— The acquisition of motor skill in young children: II The influence of specific and of general practice on motor skill *Child Develop.*, 1930, **1**, 292-297
13. HILDRETH, G. The effect of school environment on the Stanford-Binet tests of young children *27th Yrbk Nat. Soc Stud. Educ*, 1928, Pt. 1, 355-359
14. SHEPARD, J E., & BREED, F S Maturation and use in the development of an instinct *J Anim. Behav*, 1913, **3**, 274-285.
15. SPALDING, D A Instinct and acquisition *Nature*, 1875, **12**, 507
16. STRAYER, L. C. Language and growth. the relative efficacy of early and deferred vocabulary training. *Genet. Psychol. Monog.*, 1930, **8**, 209-319.

17 STUTSMAN, R Mental measurement of preschool children. Yonkers, N Y · World Book Co, 1931 Pp. x+368.

18 TERMAN, L M The measurement of intelligence Boston Houghton Mifflin, 1916. Pp. 362.

Yale University
New Haven, Connecticut

LA MATURATION ET L'APPRENTISSAGE

(Résumé)

Deux groupes, chacun de dix enfants, âgés de 24 à 36 mois, dans la "nursery school" Merrill-Palmer, ont été groupés selon l'âge chronologique, l'âge mental, le sexe, et les habiletés initiales approximées en trois opérations habiles l'action de boutonner, celle de couper avec des ciseaux, et celle de grimper On a donné douze semaines de pratique au groupe de pratique, avec de nouveaux tests à des intervalles de quinze jours Le groupe de contrôle, sans exercice spécial en attendant, a été teste à la fin de la période de douze semaines, et puis a subi quatre jours d'exercice intensif. Après cette période on a testé de nouveau le groupe de pratique et le groupe de contrôle Les différences initiales entre le groupe de contrôle et le groupe de pratique n'ont pas été sûres dans tous les tests Après les douze semaines d'exercice, le groupe de pratique a dépassé le rendement du groupe de contrôle dans tous les tests, mais une semaine d'exercice par le groupe le contrôle a suffi pour mettre les résultats du groupe de contrôle et ceux du groupe de pratique à des niveaux semblables On croit que les gains relativement rapides du groupe de contrôle signifient que des facteurs autres que l'entraînement spécifique ont contribué au développement de ces trois habiletés, des facteurs qu'on peut expliquer en partie par la maturation et en partie par l'exercice général des habiletés semblables.

HILGARD

DIE REIFUNG (MATURATION) UND DAS LERNEN

(Referat)

Zwei Gruppen, bestehend aus je zehn Kindern im Alter von 24 bis 36 Monaten, aus der Merrill-Palmer Pflegeschule (Merrill-Palmer Nursery School) (in Detroit) wurden einander gleichgestellt (equated) in Bezug auf chronologisches Alter, geistiges Alter, Geschlecht, und approximative anfangliche Fahigkeit zu drei Geschicklichkeiten. das Zuknopfen, das Schneiden mit Scheren, und das Klettern Die Versuchsgruppe wurde zwolf Wochen lang eingeubt, und jede zwei Wochen wieder gepruft Die Kontrollgruppe, die einstweilen ohne besondere Übung blieb, wurde am Ende der zwolf-Wochen-Periode gepruft, und erhielt dann vier Tage intensive Einubung Nach Verlauf dieser Periode, wurde sowohl die Versuchgruppe wie die Kontrollgruppe wieder gepruft Die anfanglichen Unterschiede zwischen Kontrollgruppe und Versuchsgruppe erwiesen sich in allen Prufungen als unzuverlassig Nach den zwolf Wochen der Einubung war

die Leistung der Versuchsgruppe der der Kontrollgruppe in allen Prufungen uberlegen Nachdem sich die Kontrollgruppe aber eine Woche lang eingeubt hatte, standen die erzielten Zahlen (scores) der Kontrollgruppe und der Versuchsgruppe auf ahnlichen Hohen. Der Meinung des Verfassers nach, weisen die relativ raschen Gewinne der Kontrollgruppe darauf hin, dass nicht nur die spezifische Einubung, sondern auch andere Einwirkungen zur Entwicklung dieser drei Geschicklichkeiten (skills) beitragen—Einwirkungen, die teilweise durch die Reifung (maturation) und teilweise durch die allgemeine Einubung in verwandten Geschicklichkeiten erklart werden konnen.

HILGARD

THE ASSOCIATION SPAN OF THE WHITE RAT*[1]

From the Animal Laboratory, Department of Psychology, Columbia University

LUCIEN H WARNER

The student of structural evolution has access to at least two important sources of data: (1) contemporary animals, and (2) extinct forms, the records and actual parts of which have by good fortune been preserved through hundreds of centuries The latter source is particularly valuable, as it yields, frequently, a somewhat continuous record of the development of a single form over an extensive period Thus the structural evolution of the locomotor apparatus of the modern horse is known in some detail Paleontology furnishes incomplete but direct evidence of the phylogenetic development of structure Information on this development from the study of contemporary animals is, perforce, indirect, since no living form is a remote ancestor of any other living form Yet, the students of functional evolution must rely chiefly upon this less satisfactory source since, unfortunately, fossil remains cannot be subjected to behavior experimentation One who is interested in the phylogenesis of the capacity for learning, for example, has no choice, but must study contemporary animals from the simplest to the most complex, in the hope that the behavior of the former may fairly be taken to represent the behavior of the remote ancestors of the latter Unless this assumption is justified, there seems little likelihood of our arriving at more than a guess as to the nature of the genesis and gradual extension of the capacity for learning.

However unsatisfactory the foundation of the study of the phylogenesis of learning, certainly the structure reared thereupon should be as soundly built as possible Strictly speaking, comparative psychology has thus far furnished remarkably few truly comparative

*Accepted for publication by C J Warden of the Editorial Board and received in the Editorial Office, December 22, 1931

[1]Part of a project carried out by the author as a National Research Fellow in the Biological Sciences under the sponsorship of Professor C J Warden of the Animal Laboratory, Department of Psychology, Columbia University

data on learning. For the most part, each animal has been studied for its own sake and with little thought of whether the results obtained would be comparable with the results obtained on other forms Most problem boxes, for example, have put a high premium on manipulatory ability and, while extremely valuable in the measurement of this particular ability, they are not well suited to a broad comparative study of learning. Even the maze, which can be employed with a far greater range of animal forms, places a high premium upon a special capacity, that of maintaining an orientation The multiple choice problem, excellent for intercomparison of higher forms, is too difficult for the learning powers of the lowest forms Comparative psychology needs further methods for the study of learning. Specifically, it needs a method (or methods) which will meet these conditions:

1 It must be adaptable to all animals, and must penalize them as little as possible for the lack of special reactive capacities or receptive capacities

2. It must provide a wide range of "difficulty" with the minimum qualitative change. In its simplest form it should be solvable by the humblest animals and yet, through only non-qualitative alteration, it should provide a form which apes cannot solve.

It should be said at once, if it need be said at all, that the writer has not produced such an ideal, universal yardstick for the measurement of learning capacity. At most, the method used in the present study can be regarded as merely a tentative step toward that goal.

The Association Span

Obviously, learning capacity cannot be studied without requiring the animal to make a response of some sort as a demonstration that it has learned The response utilized should be one which all forms can make with ease. The locomotor response is to be found in representative forms of most important animal groups. It is one of the most universal of the reactive capacities. Turning to the receptive side we find that electric shock of appropriate intensity and quality is an adequate stimulus for all animals thus far subjected to it. Furthermore, it appears that such stimulation can be so applied as to produce a locomotor response. Thus, the act of moving away from an electrified (or about-to-be-electrified) area can be used as the animal's demonstration of learning. But learning (unless mere fatigue or negative adaptation be included in the term)

must involve a second stimulus of some sort. In the method here used this stimulus serves as a warning of the impending shock. To secure truly comparative results this stimulus should with each animal be that one which it can best utilize as a warning Were light to be used arbitrarily, certain forms possessing inferior light-receptors would be handicapped The same is to be said for any other modality Naturalistic observation might give a sufficiently reliable clue as to which modality to use with a given animal It would be safer to try the several forms of stimulation to which the animal appears especially sensitive and finally to use that one with which it made the best records It is essential, also, that the warning stimulus be one which will not produce before training the locomotor response which the animal gives to the electric shock

The second requirement for the yardstick will now be considered Can the difficulty of the problem be so adjusted as to fit it for study of both the simplest and the most highly developed animals? Are many of the invertebrates capable of solving the problem of moving from an about-to-be-electrified area in response to a warning stimulus? On the basis of personal observation and of a survey of the literature an affirmative reply is hazarded, although conclusive evidence must await further experimentation The difficulty of the problem can be increased readily enough. This is accomplished simply by the insertion of a time interval between the warning stimulus and the shock. If, during the training of an animal, the shock always follows the warning stimulus immediately, the association of the two is far more readily made than if a prolonged interval always separates the two. Thus we are furnished with a quantitative measure of the animal's capacity It might be surmised that the higher mammals can synthesize in their behavior two events separated by a far greater time interval than can, let us say, a reptile, and that the maximum interval which a reptile can so span would be greater than that for an earthworm. For convenience in discussion, the term "association span" will be used to indicate that length of time which, for a given individual or species, is the maximum which can be introduced between any two stimuli still permitting their association By association of two stimuli is meant merely that the behavior in response to one of these is modified as a result of the temporal proximity of the other

It will be apparent at once that this method of measuring learning ability is not free from the "special capacity" criticism which has

already been levelled at methods now in general use. The capacity to associate events separated by an interval is special. But, among the factors governing this ability, the peculiar structure of an animal's appendages or the acuity of a given sense organ or any other single peripheral detail is reduced to a minimum. Measurement of this capacity in an animal reflects its efficiency as a unit rather than the efficiency of any one of its several reactive or receptive systems. The determination of an animal's ability to orient, to manipulate, to see, to hear, etc., throws little light on its capacity along other lines. The ability to associate over an interval is a far more general capacity for we assume that it can be utilized in connection with all of the many reactive and sensory systems of the animal.

That an animal with a long association span has a great advantage over an animal with a short one hardly needs demonstration. If there be two events, a cause and an effect, which fall within the association span of the former animal but not within that of the latter (because its span is shorter), the first animal can profit by this cause-effect relationship while the latter never will be able to do so. We should expect length of association span to correlate with biological success.

However, the significance of the association span should not be overemphasized. It represents but a single behavior characteristic and furnishes but a single yardstick, whereas many must be used in the measurement of so complex a matter as capacity for biologically appropriate behavior modification.

The association span is not to be confused with the maximum period of delay which can be introduced in delayed response experiments. The association span refers to the maximum interval which may separate two stimuli, still permitting them to be associated by an animal which is *at no time given the opportunity of responding to these stimuli when they are separated by any period less than that interval.* The interval of maximum delay is determined by delayed-response experiments is quite a different matter. To obtain the latter interval the animal is at first trained by the use of almost simultaneous presentation of the stimuli. Only then is a longer and longer period introduced between them until at last an interval is reached which results in the breakdown of the response. Pavlov and his colleagues brought about this delay by withholding the food and thus postponing the reinforcing unconditioned reflex. Hunter and other American investigators varied this procedure by mechani-

cally restraining the animal so that it could not reach the food during the delay period.

Perhaps it should also be noted that the association span is not the same thing as memory. The former relates to the temporal relationship of the components at the formation of the association. Memory relates to the durability of the association, however it may have been formed, i.e., its continued existence as a functional entity.

Experimental Literature

A survey of the literature on animal learning yields little information on the association span. The situations commonly used involve an interval which is quite brief. We cannot tell whether or not learning would have occurred had longer intervals been used. There is no adequate experimental evidence on which to base an estimate of the relation (if any) between length of the span and position in the phylogenetic scale. There is no need, then, for a survey of the literature beyond mention of certain pertinent experiments which concern the animal used in the present study, the white rat.

In Carr's (1, 2) alternation problem a single–cul-de-sac maze was used, the true pathway and the cul-de-sac being reversed on successive trials. Rats solved the problem, taking the left turn on the first run of a series, the right on the second, and so on. They learned even when trained with an interval of 35 seconds between successive runs.

A quite different learning situation has been used, with certain variations, by Carr and Freeman (3), by Wylie (4), and by Yarbrough (5). The rat, after running through a tortuous pathway, encountered a closed door, whereupon it turned about and sought another route to the goal. In a certain percentage of the trials the door was left open so that the animal's tendency to use it was not inhibited. A sound, light, or other stimulus was given a few seconds or less before the rat encountered the doorway on those trials during which the door remained closed, but not on the other trials. After considerable training, the rat turned about upon receiving the warning stimulus and did not bother to try the door. The intervals were very brief and were not measured precisely. Yarbrough, however, having taught his rats to turn when he gave them a light electric shock as a warning, proceeded to transfer this response to a sound given before the shock. The maximum interval he used

between the shock and sound was 6 seconds. Thus on the trials in which the door was closed, the rat received a sound stimulus while running down the pathway Six seconds later, if it had not heeded this warning, it received the shock Of course, if it continued despite these stimuli, it came upon the closed door. In general, the longer the interval between the stimuli, the greater the number of trials necessary for mastery of the problem. But there is no reason to suppose that 6 seconds represents the association span. In fact, the data suggest that learning might have occurred had somewhat longer intervals been used

Method

Before giving a detailed description of the apparatus and procedure, it may be best to outline briefly the situation employed. The animals were tested individually in a compartment separated into two equal halves by a low fence. An electric grid formed the floor and the animal could be shocked at any point on the floor The top of the low fence consisted of two parallel rods which served as electrodes. When the animal was given a shock, the only possible way for it to terminate this stimulation was by jumping over the fence Only one half of the floor was electrified at a time. The top of the fence was always electrified, so the animal could not take refuge there. A short preliminary training period sufficed to establish a remarkably uniform response to the shock; that of scrambling over the fence. Then the attempt was begun to establish the association. A loud buzzer was sounded for one second. After an interval (which varied in length with the different groups of animals used), the shock was given. This procedure was repeated with longer pauses between each repetition. Eventually, if the interval was not too long, the rat came to cross the fence in response to the buzzer and so to avoid the shock

General Experimental Conditions

This investigation was carried out in the writer's private laboratory, located in the country at a distance from highways, tramcars, and other sources of vibration and sound Experiments were conducted in a small basement room, originally built as a photographic darkroom. It contained a bench on which was constructed the animal-control apparatus. The stimuli-control apparatus, including relays and other devices which made some sound when operating,

was located in an apparatus room This was separated from the experimental room by 40 feet and by two walls. The sound of the relays, etc, was inaudible to the human ear from the experiment room Control tests ruled out the possibility that such sound was determining the animals' behavior A refrigeration motor located in the basement was disconnected during experiments. No other source of intermittent sound which might have affected the results was observed. The vivarium was an entirely separate building at a distance of 100 feet from the experiment room. The sound stimuli used in the latter could not have been heard in the vivarium. The rats were kept in metal cages of 1300 cu in., 2 to 4 in a cage These were hung from the ceiling An effort was made to prevent wide fluctuation in either temperature or humidity The former varied from 60 to 75 degrees Fahrenheit The percentage of saturation varied from 40 to 60 Ferns, palms, and other plants were kept in the vivarium The room was lighted by windows on all four sides, but the animals were not placed in the direct sunlight. The diet consisted of wholewheat bread and whole milk plus a daily ration of some raw vegetable such as lettuce, cabbage, beet-tops, etc The rats were fed at about 10 P.M. daily Experimentation took place between 1 and 6 P M. The regime was not interrupted on Sundays.

The white rats used were derived from stock obtained from the Wistar Institute of Anatomy and Biology, Philadelphia Only males were used in experiments. Their ages ranged from 3 to 5 months If an animal became sick during the work it was at once eliminated (there were but three such). Thus the data represent animals which were apparently healthy and normal.

Apparatus

The rats were tested in a box 9 inches long, 4¾ inches wide, and 10½ inches high (inside dimensions). It is shown in Figure 1 The walls were of wood painted flat black The floor consisted of a grid of monelmetal rods about 1/16 of an inch in diameter and each separated from the next by a gap of about 1/8 of an inch It required 49 parallel rods to form the floor 9 inches in length. The ceiling was of mathematical celluloid which transmitted light from an illumination hood mounted above it The source of this light was a Mazda flashlight bulb (2 5 volts) mounted at the center of the near side wall of the hood, 1 inch above the celluloid The

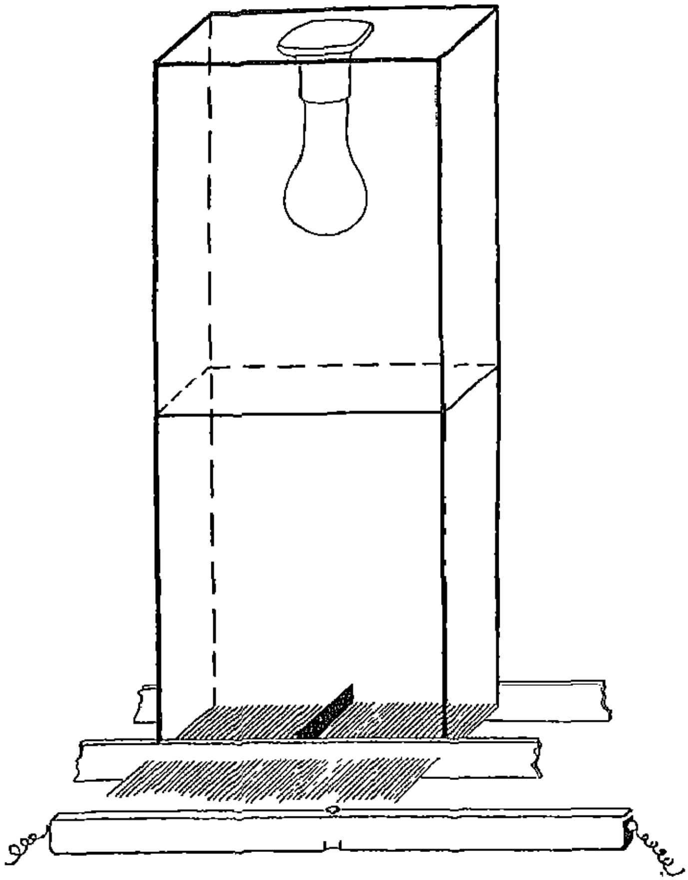


FIGURE 1
ANIMAL-CONTROL APPARATUS, SHOWING FLOOR FORMED OF RODS AND DIVIDED INTO TWO HALVES BY A LOW FENCE
(Below is shown the bar bearing electrodes.)

100-watt lamp shown in Figure 1 was not used in the present experiments. A window (not shown in the figure), 5 x 7½ inches, was cut in the near wall of the box and covered with fine wire screening painted white. Tests with the human eye showed that from the illuminated interior almost nothing of the unilluminated exterior could be descried. Nevertheless, the experimenter sitting

in the darkness without could clearly see the rat during the tests and follow its behavior in detail. The fence dividing the box into equal halves was only 7/8 of an inch in height. It was formed of a metal strip and, above this, two monelmetal rods, all mounted on the walls of the box. The strip made contact neither with the rods of the floor nor with the two which formed the top of the fence *The latter were mounted side by side, separated by 1/8 of* an inch.

In Figure 2, the box marked *Elec Stim* represents a device designed by Professor T N Jenkins of New York University, and constructed for the writer by the shops of New York University. By means of this "electrostimulator" the ordinary lighting current (110 volts, 60 cycles) can be modified within a wide range. Resistances up to several megohms can be introduced and the voltage can be varied widely It is understood that the designer will publish a detailed description of the device in the near future In the present work the resistance used was 125,000 ohms, that of step No 12 of the electrostimulator Voltages of from 40 to 60 were used, depending upon the behavior of the animal In Figure 2, the

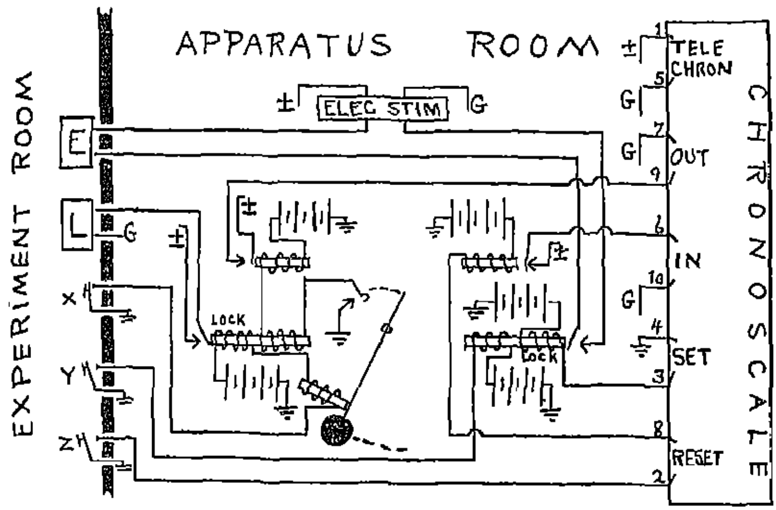


FIGURE 2
WIRING DIAGRAM
(Prepared with the assistance of Mr. J C Greene)

box marked E symbolizes the means for conveying the electric shock to the subject. Wires leading from the electrostimulator, one of them passing through a relay operated switch, ran from the apparatus room to the experiment room. There, one was attached to each electrode of the bar pictured at the bottom of Figure 1. Mounted on the bottom of this wooden bar were two copper electrodes, each extending almost one-half the length of the bar, which was itself 18 inches in length. The monelmetal rods which formed the bottom of the animal control box were mounted in two slabs of bakelite (shown with broken ends in Figure 1), and extended 4 inches outside the box as shown in the figure. Not shown is a wooden groove, the bottom of which was formed by the extending ends of the monelmetal rods. The width of this groove was such as to accommodate the bar. The copper electrodes of the bar thus made contact with the monelmetal rods. In order to stimulate the animal in the box it was necessary to adjust the bar so that one of the rods on which the animal stood made contact with one of the bar's electrodes and the other rod or rods which it touched were in contact with the other electrode. The knob with which the experimenter adjusted the bar was placed just above the half-inch gap which separated the two electrodes. Thus, to place the bar in position to shock the animal it was necessary only to move this knob opposite the body of the animal as seen through the window.

In Figure 2 the box marked L represents the warning stimulus. Specifically, the L referred to light, since a photic stimulus was used in other work on this apparatus. In all the data to be reported in this paper a sound stimulus was used, but this modifies the wiring diagram only slightly. The box marked L should here represent a buzzer. The terminals operating L (marked G and $\pm$) in this case represent the connections of a doorbell transformer. Thus, just as the lock relay operates the electric light in the diagram as drawn, so it operated the buzzer in the present experiment. The buzzer was suspended from the ceiling. It hung immediately above the center of the illumination hood and at a distance of about 2 feet above the animal's head. The buzzer gave effective stimulation, for the rats' behavior was clearly modified on the basis of its operation.

It will be noted in Figure 2 that three switches were located in the experimental room. These were knife switches mounted on the wall of the room convenient to the experimenter's left hand. They

were readily operated by touch In the preliminary training series, during which only electric stimulation was employed, the electric shock was controlled through the operation of switch *Y*, the lock magnet of the lock relay controlling the current from the electrostimulator being first disconnected Closing switch *Y* opened the relay and shut off the electric stimulation. When switch *Y* was opened the switch of the lock relay closed (being operated by gravity), and the current from the electrostimulator was transmitted to the animal The operation of the apparatus in the training series was somewhat more complicated and must be described in detail At the outset, switches *Y* and *Z* were opened Switch *X* was closed. If the wires are followed it will be seen that the current from the storage battery thereby passed through the magnet holding the pendulum and through the lock magnet of the relay which operated the buzzer To start the experiment, switch *X* was briefly opened and closed again Thus ground was removed from these two magnets with the result that the pendulum was released and the buzzer sounded (the switch of this lock relay, as in the case of the other, closed by gravity when neither magnet of the relay was in operation) The pendulum (rebuilt from chronoscope parts and with counterweight) required over 2 seconds for a double vibration Rather near the end of its outgoing swing, and exactly 1 second after its release, it operated a switch putting ground on the pick-up magnet of the lock relay, thus terminating the sounding of the buzzer. Although the pick-up magnet was thereby operated only briefly, the switch was kept open because of the lock magnet. It will be seen that in series with the pick-up magnet is a single relay which is, of course, operated by the pendulum switch simultaneously This relay started the chronoscale (a brief description of this device will be given below) Let us suppose that in this case we are using a training interval of 10 seconds, and that the chronoscale has been set accordingly. Then exactly 10 seconds after the simultaneous cessation of the buzzer stimulus and the starting of the chronoscale, the latter device (through a relay not shown in the diagram) removes ground from the lock magnet of the lock relay operating the electrostimulator This it does by closing the circuit between the terminals marked *3* and *4* The animal, providing it has not already jumped the fence, receives the electric shock This completes the presentation of one pair of stimuli To reset the device for a second presentation, switches *Y* and *Z* are used The former picks up the gravity

switch shutting off the electric shock. The lock magnet holds it. Switch Z resets the chronoscale at zero.

The *chronoscale, an entirely new device, was designed to meet* the writer's requirements by Dr. Rensis Likert who will shortly publish an article on its construction. It was built by Lockwood and Almquist of New York City. It can be used only on a frequency controlled circuit. A second chronoscale was recently built for the Animal Laboratory of Columbia University and, now that it has been standardized, one can be purchased for about $400.00. In the chronoscale is mounted a Telechron synchronous A. C. motor (terminals *1* and *5*). A clutch, operated by two pairs of powerful magnets (marked *OUT* and *IN*), connects this motor with the setting device. As here wired, both the outmagnets and the inmagnets are continuously in operation except when the current is momentarily broken through one or the other pair by the operation of the corresponding relay. Thus the clutch is held firmly against one pair of magnets until the current through that pair is broken for a moment, when it flies over to the opposite pair. The set mechanism is operated by three dials, one with 100 settings, each representing 1/100 of a second, the second with 60 1-second settings, and the third with 5 1-minute settings. Thus any setting from 1/100 of a second up to 5 minutes, in steps of 1/100 of a second, can be obtained. If the dials are arranged for 2 minutes, 39 seconds, and 98/100 *of a second, then at that interval after the clutch is thrown* in (by break of the current through the outmagnets) the connection between terminals *3* and *4* is made. The reset mechanism serves only to restore the chronoscale to zero for further trials and the reader will be spared its description.

A comment should be made on the nature of the switch operated by the pendulum. It is an enclosed mercury switch. The two terminals project into one end of a sealed glass bulb. When the bulb is so tipped that the contained mercury falls to the opposite end, the switch is opened. When tipped with the terminal end down, the mercury engulfs them, closing the circuit. This switch was fastened to the pendulum near its pivotal point. By altering the angle of its attachment it could be made to close at any desired interval after the pendulum was released. By trial and error, the correct angle for a one-second interval was obtained, the chronoscale being wired as a chronoscope for the purpose of checking this interval.

The average error of the chronoscale has been found to be re-

markably slight Data will be published by Likert It is probably true that the interval measured by the pendulum varied as much as 1 per cent It should be noted that the primary interest of the present study was in the interval separating the two stimuli, and that this interval was determined by the operation of the much more accurate chronoscale

To recapitulate briefly the above: In the training series the pair of stimuli were presented automatically, it being necessary merely to open the switch X to start the apparatus This was at once followed by the sounding of the buzzer for one second, by an interval of no stimulation (which varied in length from one group of animals to another) and then by the onset of the electrical stimulation. Current through the grid was broken manually after the animal had crossed At the end of the trial, the mechanism was reset with switches Z and Y

Procedure

When placed in the apparatus for the first time each animal was given a 5-minute period without stimulation other than that incidental to the box itself. Five 1-second stimulations of the buzzer, separated by intervals of from 15 to 20 seconds, were then given In no case did the animal cross the fence in response to this sound At the termination of this preliminary period a mild electric shock was given Only the floor of the half in which the animal happened to be was electrified, the other half remaining dead Usually, after several random efforts to avoid the shock, the rat hit upon the one successful method, that of jumping the fence. The animal generally stayed on the side to which it had jumped, but, since the current was disconnected as soon as it had crossed, recrossing entailed no punishment. In any case, at the end of 30-40 seconds whichever half upon which the rat rested was electrified and again the animal had to jump the fence to escape the stimulation After 60 such trials, given 20 per day on 3 successive days, the average rat could be depended upon to jump the fence promptly upon receiving the shock Although as many as a thousand trials were given to some rats in the course of the work, the response to the shock never failed to appear While the exact form of the response could not be called innate, certainly it appeared not to be subject to "experimental extinction"

The day following the completion of the 60 preliminary trials

with shock only, the training series was begun Each trial consisted of (1) the presentation of the sound stimulus for 1 second, (2) the training interval which varied from 1 to 30 seconds, and, finally, (3) the electric shock which continued until the rat jumped the fence Fifty trials per day were given and a careful, detailed record was made of the animal's behavior Shorthand symbols proved convenient in this connection. The time intervening between successive presentations of the pair of stimuli was not constant, but varied in an irregular fashion between certain limits.

The problem of what intervals should be introduced between successive trials was not a simple one. It was felt that these should always be greater than four times the length of the interval between the two stimuli to eliminate the probability that the two stimuli might be associated in the reverse order (the shock of one trial becoming linked with the sound of the next) With the 30-second group, the interval between trials would have to be over 2 minutes It might be argued that a between-trials interval of the same length should be used with all groups. In that case, however, the ratio between the training interval and the between-trials interval would vary widely. On the other hand, were that ratio to be kept constant for all groups at, let us say, 4 1, then, to give 50 trials to the 1-second group would require less than 5 minutes, while over 2 hours would be required to give the same number of trials to a possible 30-second group It seemed best to keep constant the number of trials per day for all groups. A compromise was therefore necessary. The basic interval between two successive trials for each of the four groups used was as follows:

For the 1-second group:	1′
For the 10-second group:	1′ 20″
For the 20-second group:	1′ 40″
For the 30-second group:	2′ 10″

An interval of 2′ 40″ was to have been used with a 40-second group, but because of the results with the other groups this one was not started. To these basic between-trials intervals was added, in each case, a certain number of seconds, from 2-12, as determined by the cast of a pair of dice Thus the between-trials interval was constantly and irregularly varied within certain limits. This prevented anticipation on the part of the animal of the exact time when each succeeding trial should commence. The same series of added sec-

onds was used for all animals Nine hundred and ninety-nine successive throws of dice were used in its preparation, there being a maximum of 1000 trials given a single animal

Immediately following the day's series in which an animal had met the criterion of 6 consecutive crossings in response to the sound stimulus, it was given 10 more trials exactly like the previous ones except that the buzzer and the grid were both disconnected. The remainder of the mechanisms and the experimenter behaved just as in the trials proper In no such case did an animal cross at the time the buzzer would have operated had it been connected. Apparently the behavior was not determined by cues other than those under investigation.

Each animal was under close observation during the entire time that it was in the apparatus In the 60 preliminary trials, during which the only stimulus given was the shock, the behavior of the animals was quite uniform. Upon first receiving the shock an animal usually jumped in an aimless effort to escape from the confinement of the apparatus Ordinarily it would merely fall upon the electrified half of the floor and again jump But occasionally it would cross the fence either directly or through a random jump Arrival on the "dead" grill usually terminated the jumping for the time being, but occasionally an animal would be so bent upon escape that it continued jumping regardless of whether it landed upon the electrified or the dead half of the floor. After each crossing, and previous to the next trial, both halves remained dead Efforts to escape usually continued throughout the 20 trials of the first day into the second day. Gradually these would be replaced by a more directed jump which would take the animal across the fence to the dead half where it would remain until the next shock was given. During the 20 trials of the third day the animals, for the most part, crossed promptly upon receiving the shock and made only infrequent efforts to escape from the apparatus

On the fourth day began the training series, consisting of 50 trials per day In each trial the sound stimulus was followed after an interval of 1, 10, 20, or 30 seconds (depending upon the group) by the electric stimulus

Results

Nature of the Data. Each animal was watched continuously Thus its behavior was observed (*a*) immediately following the pres-

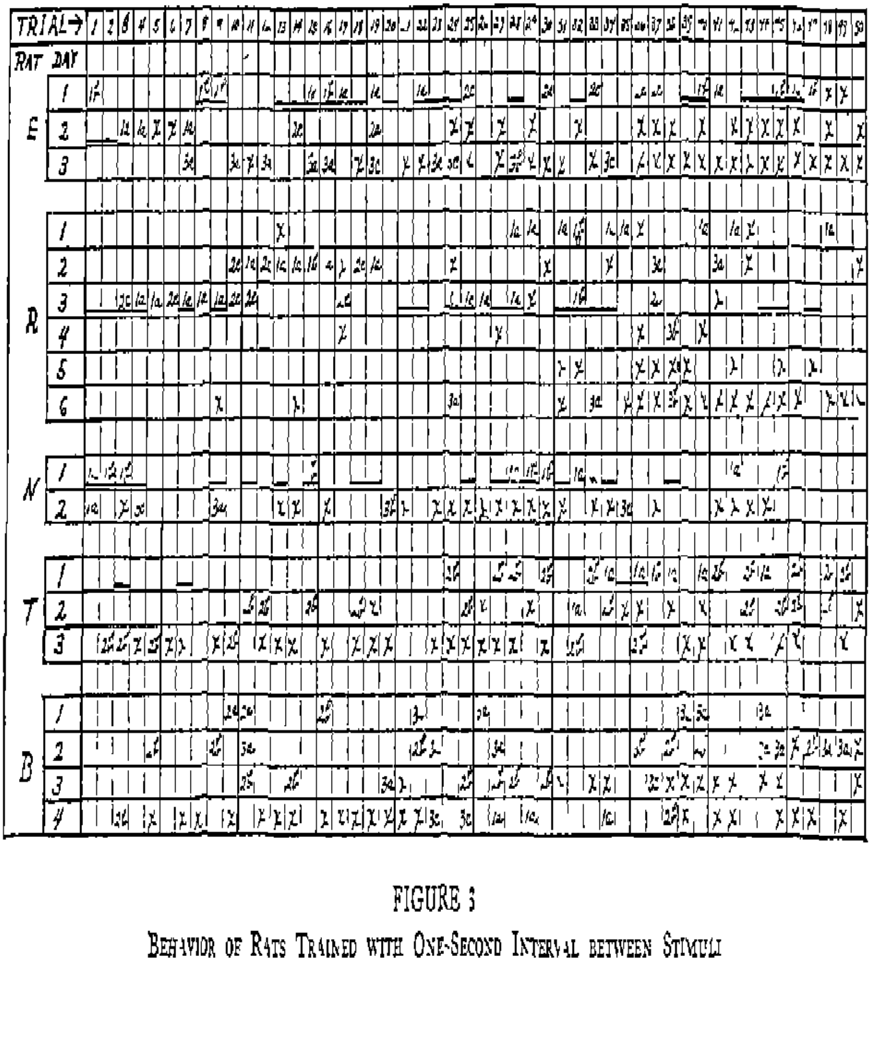


FIGURE 3

BEHAVIOR OF RATS TRAINED WITH ONE-SECOND INTERVAL BETWEEN STIMULI

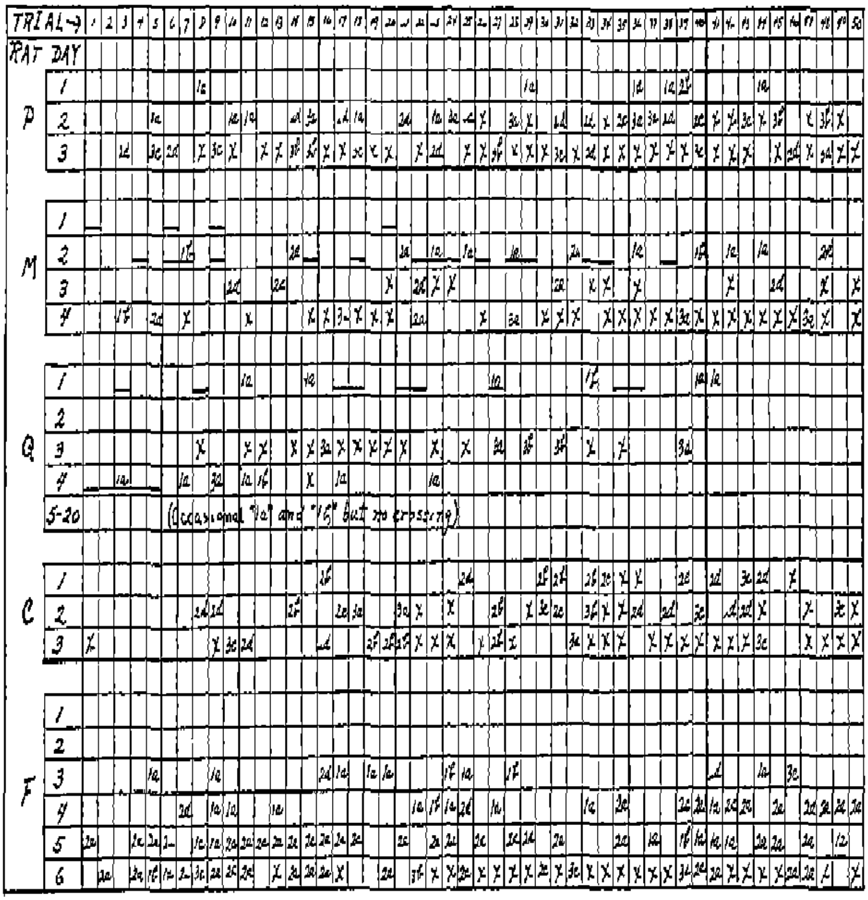


FIGURE 4

BEHAVIOR OF RATS TRAINED WITH TEN-SECOND INTERVAL BETWEEN STIMULI

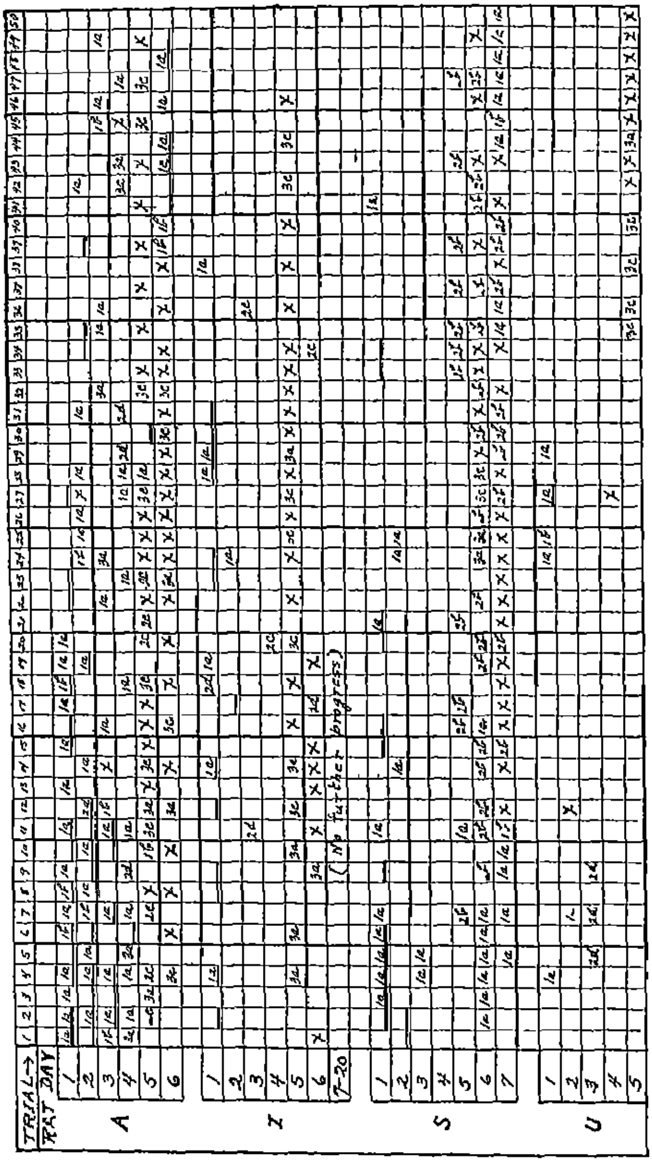


FIGURE 5
BEHAVIOR OF RATS TRAINED WITH TWENTY-SECOND INTERVAL BETWEEN STIMULI

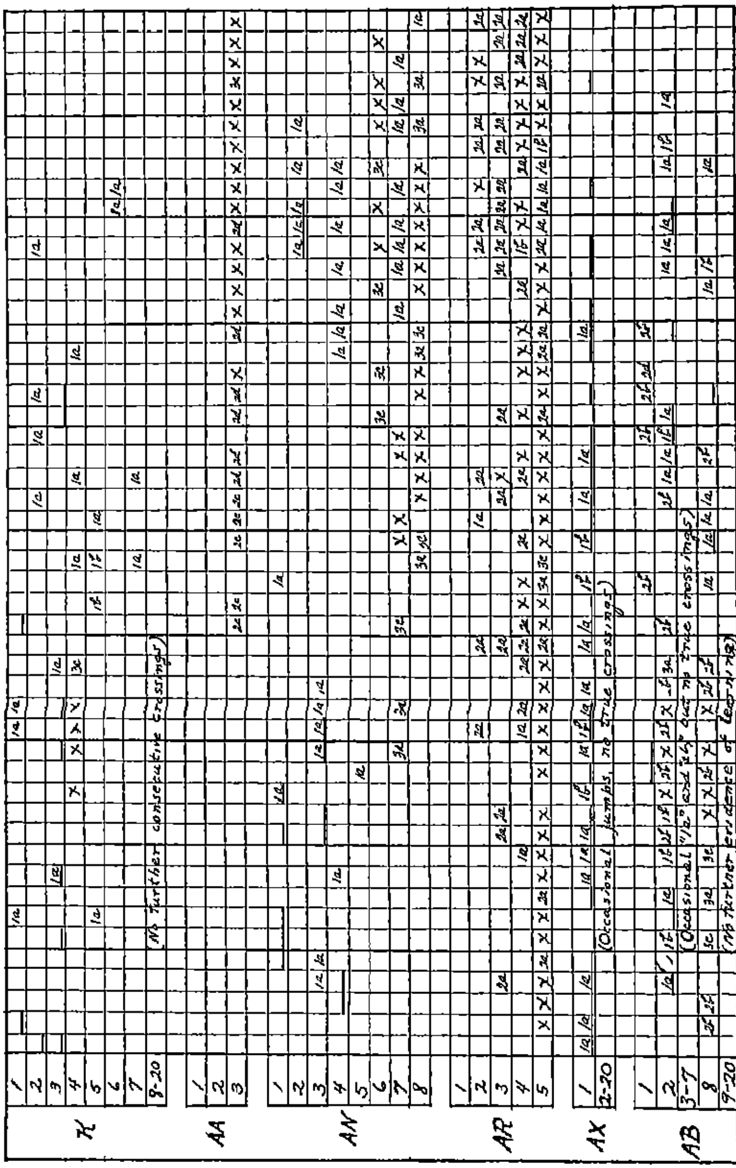


FIGURE 5 (continued)

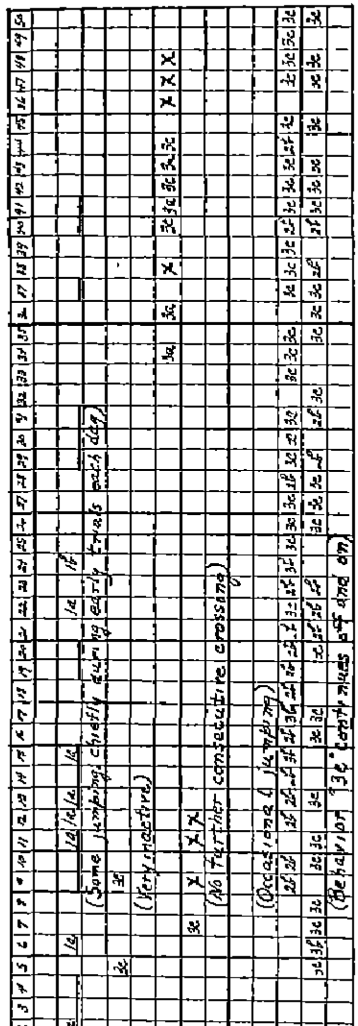


FIGURE 6

BEHAVIOR OF RATS TRAINED WITH THIRTY-SECOND INTERVAL BETWEEN STIMULI

entation of the sound stimulus and during the period preceding the shock, (*b*) immediately after the shock, and (*c*) during the period between trials A detailed report of behavior (*a*) is to be found in the accompanying tables (Figures 3, 4, 5, and 6), and this behavior is also commented upon in the verbal description of each animal which is given below The behavior made in response to, or, at any rate, following, the presentation of the sound stimulus and before the shock was found to fall rather well into four categories with subdivisions These are here given with the symbols used to represent each in the tables

1 *Abortive or misdirected behavior.* This consists of efforts to escape from the apparatus by jumping.

1*a*—indicates that after jumping the animal fell to the same side of the fence.

1*b*—indicates that the jump carried it across the fence.

2 *Undirected behavior* suggesting that the animal hears the sound.

2*a*—indicates that the animal squeaked or "chittered," having previously been silent, or that the rate of its sounds is suddenly altered (usually accelerated).

2*b*—indicates that the animal having previously uttered such sounds rather periodically suddenly ceased.

2*c*—indicates that the animal quivered or that its respiratory rate increased

2*d*—indicates that the animal ceased quivering or held its breath

2*e*—indicates that the animal ceased some other activity such as exploring or scratching

3 *Correctly directed behavior* which is, however, *uncompleted.*

3*a*—indicates that the animal turned its head toward the fence.

3*b*—indicates that it actually approached the fence and was in position to leap across

3*c*—indicates that the animal made a feint at leaping across, getting at least its forefeet and head above the fence, but then hesitated and withdrew

4 *Correctly directed and completed behavior.*

X—symbolizes the response of crossing the fence to the opposite half of the box.

The table does not indicate the behavior which followed the electric stimulation, for this was remarkably uniform. If a rat had not

crossed in response to the warning stimulus it scampered across promptly when given the shock Between trials the animals were usually quiet. But, especially early in training, there was some jumping (i.e., escape efforts) during these periods When such occurred, the symbol for the trial following that interval is underlined. When no such underline is found it may be assumed that prior to the giving of the warning stimulus the animal had not been jumping

In taking data the experimenter strove to record, not to interpret, the behavior of the animal Although a total of 11 types of behavior is distinguished in the above classification, each is sufficiently distinct from all others, and the judgment of the observer plays but a small rôle It may be objected that 1*b* and *X* are not clearly discrete The former represents an attempt to escape, a leap upwards, the rat falling on the side of the fence opposite to that from which it sprang. The symbol *X*, on the other hand, represents a crossing of the fence directly, entailing a single jump or several steps It is true that on one occasion these two forms of behavior were observed to graduate, one into the other Rat E, toward the end of the first day, made five leaps which carried it across the fence. The first was clearly an attempt to escape and was recorded as 1*b*. The last was apparently a direct crossing and recorded as *X* The others graded between these and were recorded 1*b*, although another observer might have recorded them otherwise. But this case is unique. Ordinarily, recording the rat's behavior required no such fine distinction

When no symbol at all appears in the tables for a given trial it is to be assumed either that no overt behavior was detected or (rarely) that there were mere random movements, such as scratching, which continued regardless of the buzzer. Such behavior was noted but has been suppressed from the tables as appearing to lack significance.

In a very general way, the order in which the four categories of behavior has been given corresponds to the order of their appearance. Abortive attempts to escape occur, if at all, usually in the earlier trials when, it will be noted, they frequently occur between trials as well Vague "attention" to the sound generally appears next, followed by directed response, and, finally, by successful crossing Occasionally in the notes it is found that two or more types of behavior occurred simultaneously or nearly so. In such cases the

symbol for the "higher" form of behavior appears in the tables For example, a rat upon receiving the sound stimulation might turn its head toward the fence (*3a*) and then make a feint at leaping across (*3c*) Only the latter symbol is given Or, a rat might squeak (*2a*) and, at the same time, cross the fence (*X*). Only the symbol *X* would appear It is evident, then, that the tables are a condensed version of the notes, but adequate, it is hoped, to give a true picture of the behavior changes which occurred during the course of the training series Such elaboration as is required will be found in the verbal description to be given.

Group Comparisons Of the five rats in the 1-second group, one met the criterion during the second day, two during the third, one during the fourth, and the remaining animal on the sixth day. The mean for the group falls on the third day.

The mean for the 10-second group falls on the fourth day One animal of this group failed to meet the criterion at all, although it was given a total of 1000 trials during the course of 20 days

In the 20-second group, which consisted of 10 animals, one met the criterion on the third day, two more on the fifth, one each on the sixth, seventh, and eighth days. Four of them failed although given 1000 trials The mean for this group falls at 7½ days.

Not one of the 10 animals in the 30-second group reached the criterion of 6 consecutive crossings, although each was given 1000 trials. *It is doubtful whether any amount of training would have* resulted in the required behavior. Because of the failure of the animals of this group, intervals longer than 30 seconds were not tried.

These results rather clearly indicate that with an increase in the interval separating two stimuli there is an increase in the difficulty of the learning problem involving these stimuli A comparison of the 1-second and the 10-second group suggests this trend, although the difference between the two groups is not great But with the 20-second group we find that a decidedly longer period of training is required and that a much larger proportion of animals failed entirely The 30-second interval seems to be beyond the capacity of the animals under the conditions imposed

Admittedly, the criterion of 6 consecutive crossings is purely arbitrary. However, it has been discovered that when other criteria are used (such as 5 consecutive, 3 out of 5, etc) the group relationships remain essentially unchanged

Records of the Individual Animals We turn first to the 1-second group, referring to Figure 5. The preliminary training of Rat E had not entirely eliminated its efforts to escape and during the first day it jumped frequently. Most of these jumps occurred between trials, but 13 jumps occurred during the interval after the sound and before the shock. E's jumps were vigorous and many of them carried it over the fence, but it seemed evident that such crossings were not deliberate nor in response to the buzzer. However, the crossings of Trials 48 and 49 were not jumps but slow crossings immediately following the buzzer. During the first 7 trials of the second day the animal sometimes crossed and sometimes jumped. During Trials 7 to 24 there seemed to be conflict between the two tendencies and usually the animal hesitated and then completed neither response. From this point on, however, the tendency to cross clearly predominated.

The two crossings made by Rat R on the first day may have coincided with the buzzer by chance. On the second day the buzzer frequently affected the animal. It often quivered (2*c*) upon hearing it It crossed 5 times, but not consecutively On the third day the rat reverted to jumping both after the buzzer and between trials. There was little improvement on the fourth day, the rat being very sluggish and crossing slowly even when shocked. On the fifth day, beginning with Trial 32, the animal started crossing in response to the sound and on the sixth day it well surpassed the criterion

Rat N did little but jump during the first 50 trials. On the second day, after but scant preliminary indication of learning, it started crossing and made 9 consecutive crossings in Trials 23 to 31 It crossed only 7 times during the remaining 18 trials Nevertheless, N is the record rat.

As a result of the preliminary training, Rat T almost never attempted to escape. It was a very complaining rat, making a rhythmic vocal sound most of the time. Beginning with Trial 24 the first important alteration of behavior occurred Thereafter it frequently ceased chittering (2*b*) the moment the buzzer sounded. On the second day, while crossing, it simultaneously would cease its chittering. After crossing it usually remained silent only 5 seconds or so and then resumed its rhythmic vocal complaint, continuing until the buzzer again sounded. The criterion was met on the third day.

Almost from the outset, Rat B showed a marked preference for the left-hand half of the box. When shocked in this half it would

cross to the other but would remain there only a few seconds before returning. The reason for this behavior is not known. It did not seem to affect its learning either favorably or unfavorably, for the animal crossed in response to the buzzer a dozen times on the third day and made 7 consecutive crossings on the fourth. Its solution of the problem, although of a sort entirely unforeseen by the experimenter, met the conditions which had arbitrarily been established. The rat would remain quietly on the left side until the buzzer sounded. Then it would promptly cross, remain on the right side for from 2 to 5 seconds (until after the left half had again become dead) before recrossing to its preferred side where it quietly awaited the next sounding of the buzzer.

Behavior of the rats of the 10-second group is condensed in Figure 4. Rat P paid no heed to the buzzer until Trial 14 of the second day. Preliminary evidences of attention were followed by crossings which gradually become more frequent. During the third day it either crossed or made a feint at crossing in nearly every trial.

Rat M was sluggish on the first day and spent most of the second day trying to escape. It started crossing on the third day with little preliminary indication of learning, and met the criterion on the fourth.

Animal Q appeared to make no progress during the first two days, but suddenly started crossing on the third, and just missed making 6 *consecutive crossings. It then ceased crossing almost as suddenly*, and further training yielded no signs that it would ever master the problem. During the period from the eighth to the twelfth day it was necessary to increase the shock (by degrees) from 40 to 56 volts.

Rat C was another chitterer. It early showed signs of attention to the buzzer by ceasing its chittering at the sound. It crossed 3 times during the first day, 8 during the second, and met the criterion on the third.

Rat F complained audibly, but unlike C it did not complain *chronically. Instead, it frequently started squeaking* upon hearing the buzzer and only ceased after crossing (whether before or after the 10 seconds had elapsed and the shock was given). This is represented by the symbol 2*a*. Only after many trials in which it responded only vocally to the buzzer did it start crossing. Even when it did so the movement was usually accompanied by a squeak.

Figure 5 deals with the 10 animals of the 20-second group. Rat A jumped frequently during the first four days, especially during

the earlier trials of each series On the fifth day, the jumping ceased and the animal often quivered slightly when the sound was given (2c). This preliminary sign was followed by crossing which increased in frequency until the criterion was reached. At the close of the sixth day the rat reverted to jumping

Rat I crossed very promptly when the shock was given but, during the first four days, seemed to respond not at all to the sound. The few responses indicated in the table were probably coincidences Early during the fifth day it responded to the buzzer by turning toward the fence (*3a*), but only when it was in the right-hand half of the compartment. Soon afterward it made several crossings from right to left. Beginning with Trial 26 of that day it turned when on the left side, also, and in Trials 30 to 34 it crossed in both directions Having almost met the criterion, the animal slumped and soon all signs of crossing or making any other response to the buzzer had disappeared.

Rat S was a very noisy animal, chittering most of the time. It showed no real signs of learning until the fifth day, when it suddenly ceased chittering whenever the buzzer was sounded (*2b*). About 50 trials later, crossing commenced and on the seventh day 6 consecutive crossings were recorded A spasm of escape efforts closed the day

Except for slight evidence of attention to the buzzer on the third day, Rat U showed no evidence of progress until the 35th trial of the fifth day. Then it suddenly started approaching the fence and almost crossing upon hearing the buzzer Complete crossing soon followed and the last 6 trials of the day were all crossings.

Rat K made 3 consecutive crossings on the fourth day. Aside from this it showed no signs of profiting by the warning stimulus although given 1000 trials Except for spasmodic efforts to escape during the earlier days, it was very sluggish Even in response to the electric shock it was slow and as time went on the shock had to be stepped up gradually to 60 volts in order to be effective

Rat AA was the star performer of this group. After giving no overt signs of learning for 120 trials, it suddenly started to breath rapidly when the buzzer sounded This behavior was replaced by holding of the breath and remaining tense during the period between buzzer and shock In the last 15 trials of the third day it crossed promptly 13 times

Rat AN wasted the first four days in efforts to escape Crossing

occasionally appeared on the sixth day and into the seventh Then the rat again reverted to escape efforts On the eighth day there was no jumping and the animal finally made 7 consecutive crossings

Another animal which formed the habit of squeaking in response to the buzzer was AR It started crossing on the second day and learned more slowly than most, not reaching the criterion until three days after it had started crossing

Animal AX was a total failure, although given 1000 trials. During the first 200 trials the current was increased to 55 volts which thereafter proved effective

Rat AB chittered much of the time The only signs of learning occurred on the second and eighth days when it occasionally ceased chittering at the sound of the buzzer, and even crossed But during 1000 trials it never made more than 2 consecutive crossings

We come finally to the 30-second group, represented by only two animals in the table, Figure 6, although it consisted of 10 rats The sound stimulus appeared to have a directional effect upon the behavior of only these two subjects Twice during the ninth day Rat AP seemed about to cross On the twelfth, the sound frequently affected its behavior and 3 consecutive crossings were made These were all in the same direction, the animal having recrossed each time during the between-trials interim Three crossings occurred on the thirteenth day, but during the remaining six days no certain signs of directional response to the buzzer were detected

Rat AL made no crossings at all, but for several days beginning with the eighth it seemed on the verge of mastering the problem. During the first week it had made no certain response to the sound But on the eighth day its vocal complaints, which had been more or less continuous from the beginning, were suddenly stifled almost every time the buzzer sounded, and on many occasions this behavior was accompanied by a turning toward the fence Often one paw would be placed upon the nearer of the two rods forming the top of the fence and the animal seemed about to cross These responses always occurred immediately after the warning. After a few seconds in this "preparatory" position, the rat would retire to a corner and when, 30 seconds after the buzzer, the shock was given, it was not at all ready to cross and scrambled over the fence as best it could. This behavior continued occasionally throughout the remainder of the animal's training

None of the eight remaining rats of this group gave such definite

signs of promise and detailed data need not be reported here, although these may be referred to in a later article Each rat was given the usual 1000 trials. The rats of this group, and to a lesser degree those of the 20-second group, were never so alert as those of the groups with shorter intervals Especially after the first two or three hundred trials were they inclined to relapse into what appeared to be a drowsy state—interrupted only by the shock.

Discussion

These experiments show that the matter of determining the association span of the rat is not simple Certainly, it cannot be measured exactly unless many qualifications are stipulated. The criterion of learning is but one of these. Many rats with a 1-second interval can finally make 15 or more consecutive crossings, but few rats, if any, can exceed this number using a 10-second interval Thus if 15 consecutive crossings were adopted as the criterion we should have to place the association span for the typical rat at less than 10 seconds But more than half of the 20-second group attained 6 consecutive crossings, so if that were our criterion we should have to place the span at least as high as 20 seconds. However, even a most lenient criterion would rule out the 30-second group.

The question arises as to why it is that the number of consecutive crossings before a lapse occurs is always limited. The answer is easily found and relates to the cause of the lapses. In the present experimental situation, the act of making the required response (crossing in response to the sound and before the shock) automatically removed the possibility of the animal's receiving the reenforcing stimulus. Thus, the instant association is accomplished, the dissociation process begins. Because of this it appears unreasonable, in the present experiment, to demand a very high criterion. That of 6 consecutive crossings is adequate to demonstrate that the association has occurred According to that criterion, the rat's span, under the conditions of the experiment, is at least 20 seconds in length and is less than 30 seconds.

Among these "conditions of the experiment," which have an unknown bearing upon the results, several will be mentioned. Firstly, only male rats were used Secondly, only a single pair of stimuli was employed. Whether or not the span would have been the same had light been substituted for sound, for example, is not known. Preliminary work had shown that learning was slower when light

had been used as a warning stimulus Furthermore, there was no study of the effect of varying the intensity of the stimuli The intensity of the sound was determined only in an effort to have it loud enough to be effective, without being so loud as too seriously to startle the animals. Similarly, the electric shock was strong enough to force the animals to cross promptly and yet was the minimum intensity which could be depended upon to have this result. Another factor of unknown influence is the effect of distribution of effort. Only the one procedure, 50 trials a day, was used Similarly, the relation of the length of time between successive trials is unknown.

Only one more of these factors will be mentioned here: the nature of the response. The crossing of the fence was a gross bodily movement of which any normal rat is capable It was not required that the animal should cross the fence in any particular manner, but merely that, upon receiving the warning stimulus, it should quit the side upon which it happened to be and get itself entirely over the fence (Occasionally the tail, or a part of it, remained on the quitted side, but a concession was made on this point. If only it remained the rat was considered to have crossed and it should be noted here that the rats did not appear sensitive to the shock when only the tail touched the electrodes) Had a more complicated or difficult response or system of responses been required of the animals, it is quite probable that they could not have associated stimuli separated by so long an interval

A number of miscellaneous observations arise from examination of the data The "warming up" effect is frequently demonstrated The number of crossings made during the first dozen or 20 trials of each day is decidedly lower, usually, than that made during the remainder of each day's 50 trials.

Certain rats, as for example Rat I, learned first to cross in one direction, and only later learned to cross in the other. Rat A, on the other hand, after crossing frequently in both directions, for a time ceased crossing from left to right and only crossed from right to left (end of fifth day). Usually, however, a rat made no distinction between the two halves and crossed, if at all, from either indiscriminately Rat B always distinguished between the two halves as has been already noted

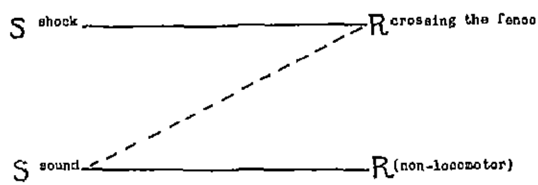
The conflict between jumping and crossing has been mentioned above in connection with Rat E, but it was also noted in other cases. This conflict was of a sort not easily recorded in an objective manner

and yet it often was apparent to the observer It occurred as a transitional stage. After the usual period of efforts to escape, occurring as much between trials as during the interval following the buzzer, came often a stage during which the animal rested quietly until the buzzer sounded and then started jumping Following this, and before the occurrence of frequent crossing, the animal upon hearing the buzzer sometimes would brace for a jump and then apparently inhibit it and perhaps approach the fence instead Occasionally an animal would look upward and then lower its head and appear to look across the fence but make no other movement in either direction

The problem of abrupt learning deserves attention. Those of the Gestalt school have interpreted learning, in certain cases, as being a flash of insight rather than a gradual process With many of the rats tested, a long period during which gradual progress could be observed preceded mastery of the problem. First the jumping would be abandoned. Then, perhaps, behavior of the "2" variety (*2a, 2b, 2c, 2d, 2e*) would appear, indicating that the buzzer was beginning to play a part in the determination of the animal's behavior Then would follow responses of the "3" variety. The subject would turn toward the fence and almost, but not quite, leap across upon hearing the sound Next it would actually cross occasionally, and, finally, these crossings would become frequent and consecutive for a number of trials. There were, however, rats which learned suddenly, if we are to judge from the tables. Rat N is the outstanding example. Its first crossing was not preceded by anything but jumping, and it made 9 consecutive crossings with no preliminary behavior being recorded except one *3a*, one *3b*, and one *3c*. In the opinion of the experimenter, nevertheless, no learning, not even that of Rat N, was sudden or could by any stretch of the imagination be described as indicating a flash of insight As stated above, the tables represent a condensed version of the notes and include only the most clear and obvious observations. Certain forms of behavior could not be observed so reliably and were, therefore, omitted from the tables One of these was the "bracing" behavior occasionally seen during the trials before an animal started crossing. It was a very slight movement which, if exaggerated, could be described as hunching the back and drawing the feet close to the body. Usually at this time the head was lowered and the nose almost touched the rods. An anthropomorphic observer would have said that the sound had made

the animal fear that a shock was about to follow Another point not brought out in the tables is the effect of training upon the speed with which the animal crossed when given the shock. This speed was not measured and thus too much weight cannot be given to the observation Nevertheless, it often seemed evident that the speed increased considerably, and this increase usually occurred during the trials preceding the occurrence of crossing in response to the sound Perhaps the sound stimulus came to have an accelerating effect upon the response to the shock Taking these various symptoms altogether, the experimenter could frequently predict successfully that the animal was nearing mastery of the problem In other words, learning progressed gradually rather than suddenly, and this gradual process could be detected even though there were noted only such gross movements as could be observed by the experimenter visually. Unquestionably, there were other evidences of incipient learning which entirely escaped observation.

The learning problem used in this experiment was quite simple as compared with those usually employed with mammals It may be inquired whether or not it can be described in terms of the conditioned response Superficially it simulates a conditioned response and one might be tempted to diagram the behavior modification thus



It is evident that if this diagram correctly represents the process, the response of crossing must be the *same* response whether made as a result of the shock or of the sound stimulus In the opinion of the experimenter this was never the case Crossing in response to the shock and crossing in response to the sound were different behavior patterns They had little in common except that they were made by the same animal, using the same four legs, etc , and except that both responses succeeded in transporting that animal from one side of the fence to the other Unfortunately, motion picture data in support of this opinion cannot be offered. The response to the sound was a

smooth, unhurried, though fast, movement It was a clean, neat jumping of the fence following the sound stimulus without hesitation and yet without haste. Little energy was wasted, the rat raising itself just high enough to clear the low fence This differed radically from the crossing in response to the shock The latter was usually jerky, sometimes almost hesitant, involving one or more very rapid removals of the feet from the rods before the rat scrambled ungracefully over the fence. Much energy was wasted and the movement suggested rather frantic, but somewhat inefficient, haste Response to the sound involved a steady push of the feet which threw the animal over the fence The response to the shock was a jerky, hasty hop. The effect of the two responses was the same but the neuromuscular mechanism employed was certainly not identical The statement that the response originally aroused by the shock was, after training, aroused by the sound is fallacious. The training certainly modified the animals' behavior, but the neurological basis of that modification was surely not that indicated by the conditioned response diagram.

It is not contended that it is impossible to reduce learning to terms of the conditioned response But, if this reduction is possible, it must involve more delicate methodology than here employed. The learning herein observed is not reduceable to one or two conditioned responses. Perhaps it is reduceable to a multitude of them, superimposed and interacting, and detectable only by means thus far undeveloped.

Conclusion

It is evident that when the interval between the two stimuli was twenty seconds or less, the response to the sound stimulus was clearly altered, there becoming attached to it a behavior pattern which yielded the same end-result (that of getting the animal across the fence) as that produced by the shock We can conclude that with most white rats the response to one stimulus can be altered as a result of the regular presentation of a second stimulus one, ten, and even twenty seconds later. We can conclude further that, in the present situation, the response is not altered when an interval of thirty seconds separates the two stimuli.

References

1 Carr, H A. The alternation problem *J Anim Behav*, 1917, 7, 365-384.

2 ———— Length of time interval in successive association. *Psychol Rev*, 1919, **26**, 335-353

3 CARR, H A, & FREEMAN, A. S Time relationships in the formation of associations *Psychol Rev*, 1919, **26**, 465-473.

4 WYLIE, H H An experimental study of the transfer of response in the white rat *Behav Monog*, 1919, **3**, No 16 Pp iii+66

5 YARBROUGH, J U The influence of the time interval upon the rate of learning in the white rat. *Psychol Monog*, 1921, **30**, No 135 Pp 52

Columbia University
New York City

LA PLUS GRANDE DURÉE DE L'ASSOCIATION CHEZ LE RAT BLANC

(Résumé)

Le terme "association span" (la plus grande durée de l'association) est defini comme "la durée qui, pour un espèce ou individu donné, est le maximum qu'on peut introduire entre n'importe quels deux stimuli sans la perte de leur association, l'association de deux stimuli veut dire seulement que le comportement dans la réponse à l'une de ceux-ci est modifié comme résultat de la proximité temporelle de l'autre" On a testé les rats individuellement dans un compartiment separé en deux moitiés égales par une barrière peu élevée Le plancher s'est composé d'une grille électrisée et l'animal a pu recevoir un choc n'importe où sur le plancher La partie supérieure de la barrière s'est composé de deux petits bâtons parallèles servant d'électrodes Quand l'animal a reçu un choc, il n'a pu terminer cette stimulation que par un saut par-dessus de la barrière On n'a électrisé que la moitie du plancher à la fois. La partie supérieure de la barrière a été toujours électrisée, en sorte que l'animal n'a pu s'y réfugier Une petite période préliminaire d'entraînement a suffi pour établir une réponse très uniforme au choc celle de sauter par-dessus la barrière Ensuite on a commencé à faire etablir l'association On a fait sonner un timbre electrique retentissant pendant une seconde Après un intervalle (lequel a varie en longueur avec les divers groupes des animaux employés) on a administré le choc et l'animal a sauté par-dessus la barrière Ce processus a été répété avec de plus longues durées entre chaque répétition Enfin, si l'intervalle employé n'a pas été trop long, le rat est venu sauter par-dessus la barrière en réponse au timbre et ainsi a évité le choc On a adopté comme critère de l'apprentissage six sauts consécutifs en réponse au son et avant le choc Les stimuli employés ont été définis assez exactement et on a gouverne le temps très précisément par l'emploi d'un mécanisme électrique compliqué gouverné par un moteur synchrone à courant alterné Les résultats indiquent que l'accroissement de la longueur de l'intervalle qui sépare les stimuli augmente la difficulté de l'association, mesurés ou par le pourcentage des animaux qui sont arrives au critère ou par la durée de l'entraînement exigé Le critère a été atteint par tous les animaux entraînés avec un intervalle d'une seconde, par quatre des cinq entraînes avec un intervalle de dix secondes, par six des dix entraînés avec un intervalle de vingt secondes Pas un des rats entraînés

avec un intervalle de trente secondes n'a atteint le critère ni même a fait de grands progrès Dans les conditions imposées, la durée de l'association chez le rat semble être moins d'une demi-minute L'expérimentateur note le fait que le saut par-dessus la barrière en réponse au son et le saut en réponse au choc ont paru être tout à fait dissemblables qualitativement

WARNER

DIE ASSOZIATIONSSPANNE DER WEISSEN RATTEN

(Referat)

Der Ausdruck Assoziationsspanne wird definiert als "jener Zeitabschnitt, der für ein gegebenes Individuum oder eine Art, das Maximum darstellt, das zwischen irgend zwei Reizen eingeschaltet werden kann, und trotzdem noch deren Assoziation zulasst, unter Assoziation zweier Reize versteht man lediglich, dass das Verhalten in der Reaktion auf einen derselben verandert wurde, als Ergebnis der zeitlichen Nachbarschaft des andern" Die Ratten wurden einzeln in einem Abteil gepruft, das durch einen niederen Zaun in zwei gleiche Halften geteilt war Eine elektrisierbare Platte bildete den Boden, und dem Tier konnte in jedem beliebigen Punkt des Bodens ein elektrischer Schlag versetzt werden Der obere Rand des niederen Zaunes bestand aus zwei parallelen Stangen, die als Elektroden dienten Wenn dem Tier ein Schlag versetzt wurde, konnte es der Reizung nur durch einen Sprung uber den Zaun entfliehen Zu einer bestimmten Zeit wurde nur die eine Halfte des Bodens elektrisiert Der obere Rand des Zaunes war immer elektrisiert, sodass das Tier dort keine Zuflucht fand. Es genugte eine kurze Übungszeit, um eine auffallende Gleichformigkeit in der Reaktion auf den Schlag zu erzeugen namlich Hast uber den Zaun Dann bemuhte man sich, eine Assoziation herzustellen Ein lauter Summer gab ein Signal, das eine Sekunde dauerte Nach einem Intervall (das bei den verschiedenen Tieren hinsichtlich der Lange variierte) wurde der Schlag versetzt, und das Tier sprang uber den Zaun Dieses Verfahren wurde mit langen Pausen zwischen je zwei Versuchen wiederholt. Schliesslich, wenn das benutzte Intervall nicht zu lange war, begann die Ratte als Reaktion auf den Summer uber den Zaun zu springen und vermied den Schlag Sechs nacheinanderfolgende Sprunge als Reaktion auf den Ton und vor dem Schlag wurden als Kriterium des Lernens aufgefasst Die Reize wurden ziemlich exakt definiert und wurden zeitlich genau abgemessen, indem man einen sorgfaltig ausgearbeiteten elektrischen Zeitmesser einschaltete der durch einen synchronomen A C Motoren reguliert wurde. Die Ergebnisse zeigen an, dass eine Verlangerung des Intervalls, das die Reize von einander abgrenzte, die Schwierigkeit der Verassoziierung vergrossert, die entweder durch die Prozentzahl der das Kriterium erreichenden Tiere, oder durch die notwendige Übungszeit gemessen wurde Das Kriterium wurde von allen erreicht, die mit einem Zeitintervall von einer Sekunde trainiert wurden, aber nur von vieren der funf deren Übungsintervall zehn Sekunden betrug, und von sechs der zehn, die einem zwanzig Sekunden Intervall unterworfen wurden Nicht eine einzige der zwanzig Ratten, deren Intervall dreissig Sekunden betrug, erreichte des Kriterium, oder wiesen irgend welchen Fortschritt auf in dieser Hinsicht Unter den auferlegten Bedingungen scheint die Assoziationsspanne weniger als eine halbe Minute zu sein Der Versuchsleiter stellte die Tatsache fest, dass das Überspringen des Zaunes als Reaktion auf den Ton und dasjenige als Reaktion auf den elektrischen qualitativ ganz ungleich waren.

WARNER

AN EXPERIMENTAL SEARCH FOR THE "CONDITIONED RESPONSE"*[1]

From the Animal Laboratory, Department of Psychology, Columbia University

LUCIEN H WARNER

While conducting experiments on the association span of the white rat (1) the writer frequently observed behavior modification of a sort which one is tempted to describe as a simple conditioned response Rats proved able to profit by the frequent repetition of a warning signal (loud buzzer) which was always given one second before an electric shock transmitted to the animal through the floor on which it stood. Before training, the animals paid no apparent heed to the buzzer After considerable training they would leap over a low fence and off the about-to-be-electrified section of the floor in response to the sound and before the shock was given It might be supposed that this simple case of learning would admirably fit the conditioned response formula The fence-jumping response was originally given *only* to the shock. After training, the sound seemed to arouse this response Thus one is tempted to term the sound a "substitute stimulus" But if this description is accurate the crossing of the fence in response to the sound *must be identical* with the crossing in response to the shock. The writer was persuaded that this was not true. To quote·

> "The response to the sound was a smooth unhurried though fast movement. It was a clean, neat jumping of the fence following the sound stimulus without hesitation and yet without haste Little energy was wasted, the rat raising itself just high enough to clear the low fence This differed radically from the crossing in response to the shock The latter was usually jerky, sometimes almost hesitant, involving one or more very rapid removals of the feet from the rods before the rat

*Accepted for publication by C J Warden of the Editorial Board and received in the Editorial Office, December 22, 1931

[1]Part of a project carried out by the author as a National Research Fellow in the Biological Sciences under the sponsorship of Professor C J Warden of the Animal Laboratory, Department of Psychology, Columbia University

> scrambled ungracefully over the fence Much energy was wasted and the movement suggested rather frantic but somewhat inefficient haste Response to the sound involved a steady push of the feet which threw the animal over the fence The response to the shock was a jerky, hasty hop The effect of the two responses was the same but the neuromuscular mechanism employed was certainly not identical"

Nevertheless, this opinion of the observer carries no great weight since the two responses had some elements in common, and it might well be maintained that another observer would have given quite a different interpretation of the behavior

It seems worth while, then, to submit a report of other experiments conducted in an effort to devise a situation which would yield behavior modification which could truly be described as a single "conditioned response"

Situation I—The Box

The rat was placed in a box 9 inches long, 4¼ wide, and 10½ high. The floor of this box consisted of an electric grid. The ceiling was formed of mathematical celluloid, above which was mounted a 100-watt Mazda lamp. Also above it was mounted a 2.5-volt lamp operated by a dry cell. This box was that described and pictured in the article cited above (1) except that there was no fence dividing the floor into two halves The source and nature of the electric current supplied to the grid has also been described, as has the timing mechanism for controlling the presentation of stimuli

It was planned that an electric shock of one second's duration should be the primary stimulus The response to be conditioned was whatever one the rat normally made to the shock The secondary, or to-be-substituted stimulus, was to be the light furnished by the 100-watt lamp which was to flash on for just one second. The 2.5-volt lamp remained lighted throughout to permit the experimenter to observe the animal

1 First, four animals were trained separately under the following procedure. The light flashed on for one second, then followed a one-second interval, then the grid was electrified for one second This sequence constituted a "trial" The interval between trials was 18, 20, or 22 seconds, which of these being determined by chance. It was supposed that this was an ideal situation for conditioning. The light and the shock appeared to be the dominating

stimuli because of their intensity and because they were given against a uniform background, i e., there were no distraction or confusion stimuli. After a sufficient number of trials the response given originally to the shock might be elicited by the presentation of the light alone, the shock being entirely omitted Every 20th trial was such a "control" presentation

Results There was no response uniformly made to the shock Sometimes the rats jumped, sometimes they ran about, sometimes they merely lifted their feet gingerly. Thus there was no clear-cut response to be conditioned However, before the end of the 100 trials which were given to each animal in a single series the light did seem to play a part in the situation It often stirred the rats to general activity, whereas during the period between trials they were often inactive

2 For the next four rats (no animal was used in more than a single experiment) the procedure was varied in an effort to secure a single specific response to the shock The duration of the shock was reduced to ½ second and a preliminary series of 100 trials with the shock only was given the day before the presentation of light followed by shock.

Results By the end of the preliminary 100 trials the response of each animal had become more stereotyped Exploration and escape efforts had almost ceased When shocked, they no longer jumped high with head and forepaws up in an attempt to get through the celluloid ceiling Instead, they made a series of two or three rapid hops with head lowered so that the nose almost touched the grid The hopping ceased with the termination of the shock, but their respiration, accelerated by the shock, returned to normal more gradually. Apparently a very specific response had been discovered but the results of the efforts to condition this response were disappointing. The animals were given 1000 trials, 100 per day, with the light preceding the shock by one second At no time, either in the regular or in the control trials, did the rats give the hopping reaction in response to the light On the contrary, when the light was given they usually lowered their heads, and crouched tensely. True, one rat breathed rapidly in response to the light just as it originally had done only when the shock was given But the other rats usually did the reverse, holding their breath for the light, breathing rapidly after the shock

3 Four rats were then trained exactly as were those in the pre-

ceding experiment except that a loud buzzer was substituted for the light.

Results These animals began to respond to the sound during the first 50 trials, whereas those of the preceding experiment paid no evident attention to the light until after 120 trials had been given. Otherwise there was little difference. One of these rats frequently squeaked for the shock and sometimes in apparent response to the sound and before the shock Otherwise no sign of conditioning was noted.

4 Four rats were trained as in the preceding experiment except that the sound and the shock were contiguous in time instead of being separated by a one-second interval. Again there were no signs of conditioning although 500 trials were given to three of the rats and 1200 to the fourth, which at one time showed vague promise of becoming conditioned.

Having failed to transfer the response from the shock to the to-be-substituted stimulus with any of these 16 rats, we abandoned this method.

Situation II—The Suspended Cage

It was thought that, even though no gross motor response had been conditioned, perhaps some less easily detected response had been or might be. The breathing rate, in several cases, seemed to have been obviously altered by the training. An effort was therefore made to construct a device which would detect and record movements which might escape the experimenter's unaided observation.

A small box was constructed of three pieces of mathematical celluloid and two sheets of copper. It is shown in Figure 1 The two curved pieces, shown shaded in the figure, were copper The two shield-shaped sides and the square lid were celluloid. The copper sheets were cut with tabs which fitted into slits cut in the celluloid and were bent over. The lid was hinged and could be opened for the insertion and removal of the animal or fastened shut to prevent its escape. This cage was suspended (see Figure 2) from the four upper corners The two copper sheets served as electrodes Because of the shape of the cage, the smoothness of its walls, and the force of gravity, it was impossible for a rat to avoid making contact with both copper sheets continuously. It could not avoid contact with the two electrodes The gap between the two copper sheets (at the bottom) was ⅜ of an inch. This was too narrow to afford the

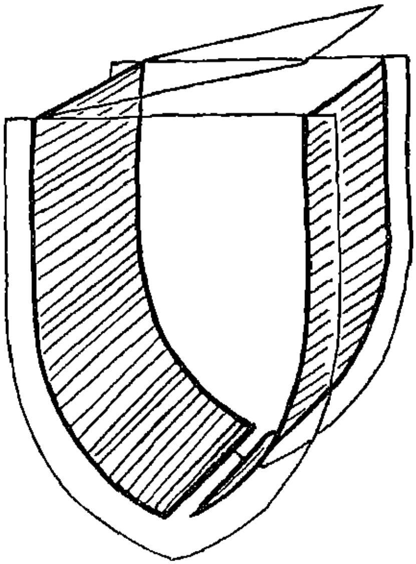


FIGURE 1
THE SUSPENDED CAGE

rat a means of escape and yet wide enough to prevent the rat's fouling the electrodes and, through the resultant short-circuit, escaping the shock Other dimensions of the cage were as follows· width of copper sheets, 5 inches, distance between the two copper sheets at top, 4¾ inches, maximum width of celluloid sides (shield-shaped), 6 inches, height of sides, 8 inches, distance between top of cage and point at which copper sheets begin to curve inward, 4 inches.

As shown in Figure 2, this cage was suspended inside a box The four wires supporting the cage converged at a point 10 inches above it and the single wire passed through a small opening in the top of the box and was attached to the center of a rubber diaphragm, 4 inches in diameter. This diaphragm was stretched across the larger end of a funnel which was firmly mounted above the box From the small end of the funnel there ran a rubber tube which was attached, at its other end, to a Marey tambour in another room The tambour recorded on a kymograph which is indicated in Figure 3

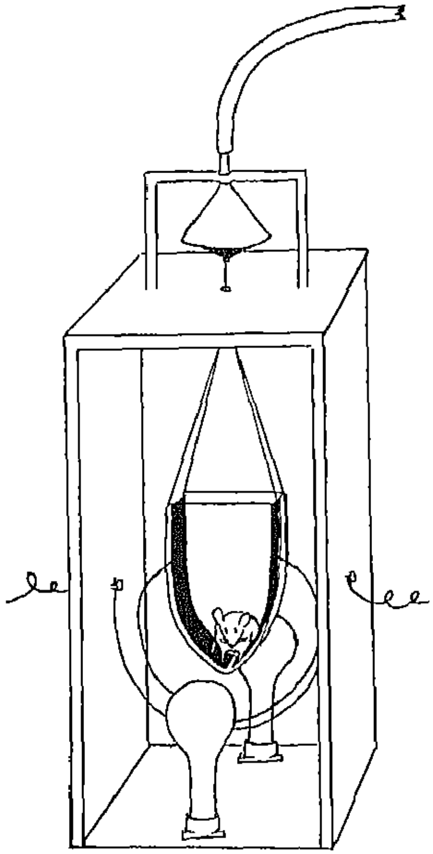


FIGURE 2
THE SUSPENDED CAGE IN POSITION

Thus, even a slight movement made by a rat in the suspended cage was recorded on the smoked paper of the kymograph. The box in which the cage hung was constructed of wood and lined with heavy grey cardboard It was 26 inches in height and 1 foot square (inside dimensions). When suspended and containing a rat of average weight, the top of the cage was about a foot below the ceiling of the box Mounted on the floor of the box, as shown in Figure 2, were two 100-watt Mazda lamps, so placed that the light from them was

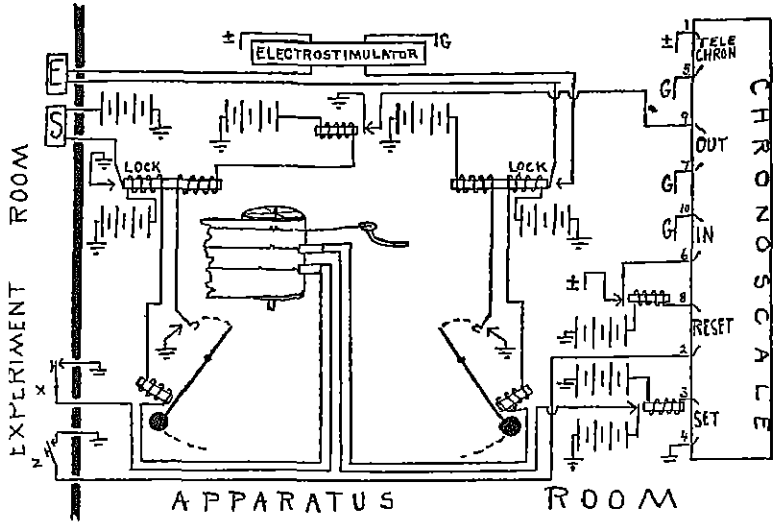


FIGURE 3
WIRING DIAGRAM

thrown through the celluloid sides of the suspended cage and directly on the rat. The centers of these lamps were separated by a distance of 9½ inches. Attached to each of the two copper sides of the suspended cage was a wire. These led to binding posts on the wall of the box and thence to the electrostimulator. The wires from the binding posts to the suspended cage were arranged in such a manner that they interfered hardly at all with any vertical movement of the cage and yet served to counteract any tendency for the cage to rotate on its vertical axis. The latter function was important since such rotation would have placed the opaque copper side of the cage between the lamps and the rat and it would also have interfered with the experimenter's observation of the animal. The arrangement of these two wires is clearly shown in Figure 2. In this figure the front wall of the box has been removed (as it could be for insertion and removal of the rat). In this front wall was cut a small window (6 inches in width and 1 inch high) through which the animal could be watched. The window was covered with wire screening painted white. Since the illumination was entirely from within, the animal was clearly visible to the experimenter, while the latter, being in nearly complete darkness, could hardly have been

seen. (This technique of observation was followed with appropriate modifications in all the work reported in this article) Not shown in Figure 2 is the 2.5-volt pilot light which was mounted on the ceiling of the box in the middle of the front side. This remained lighted continuously, whereas the two 100-watt lamps only flashed on occasionally as stimuli. Suspended immediately above the box was a buzzer which in some experiments replaced the 100-watt lamps as source of the secondary stimulus.

Since the wiring in this set-up differs in some respect from the arrangement used previously (1), it seems best to publish another diagram (Figure 3). Located in the experiment room (in which was the box with suspended cage) were two switches, *X* and *Z*. The former set the apparatus into operation for a given trial and the latter reset it in preparation for the succeeding trial. *E*, in the diagram, represents the shock stimulus, i.e., the two copper electrodes forming walls of the suspended cage. *S* represents the secondary stimulus, the buzzer in certain experiments and the two 100-watt lamps in others. It will be noted that the opening of key *X* gives the secondary stimulus and at the same time releases a pendulum which, after an adjustable interval, terminates this stimulation. At the same time that it terminates it, it also starts the chronoscale, and this device, again after a predetermined interval, gives the primary stimulus (shock) and simultaneously releases the second pendulum which terminates the shock one-half second later. The moment at which each of these stimuli is given is recorded on the kymograph record by electric markers appropriately wired. In the small bit of record shown in Figure 3 the lower line represents the secondary stimulus (which is always given first), the middle line represents the electric shock, and the upper line is that made by the tambour operated by the rat's movements in the suspended cage.

The suspended cage and associated recording devices as originally constructed did not record the animal's breathing entirely satisfactorily. It was both too sensitive and insufficiently sensitive. When the rat was quiescent and breathing normally respiration was not recorded at all, the kymograph record being a straight line. Although the suspended cage was as light as it could well be built and still be rigidly constructed and bear two electrodes, it nevertheless possessed sufficient inertia to dampen the effect of the very slight respiratory movement. Labored breathing, however, made a faint record when the animal was otherwise quiet. Obviously, a more

sensitive arrangement would be required to differentiate clearly changes in breathing. On the other hand, the apparatus was so sensitive that the violent movement of the animal in response to shock quite obscured any other movement made at the same time or for 2 or 3 seconds thereafter Even though the rat remained quiet after a single violent movement, the system remained in vibration for a brief period Subsequent alterations of the apparatus may overcome these defects. But, even as constructed, the device was of service if only because it provided an objective record of the rat's movements, free from any interpretation (or misinterpretation) by the experimenter

A comparison of the experimenter's notes of the rats' behavior with the kymographic record of that behavior was helpful in several ways. It enabled one to determine rather exactly the moment when each of the rat's observed movements was made and to relate these movements to the time of presentation of the stimuli (since this was also recorded on the kymograph). And, as mentioned above, it enabled the observer to secure an objective check upon his observations as to the nature and the violence of the several movements. As a matter of fact the kymographic record added little in the way of new data since it brought to light little not detected visually and since it uniformly verified the visual observations. It is hoped, however, that in later work the method can be so refined as to yield data on the animals' behavior which cannot be secured by unaided observation.

1. At first, light was used as the secondary stimulus. The two 100-watt lamps would flash on for one-half a second Then would follow an interval of one second which was terminated by the shock stimulus which lasted for one-half of a second. The interval between consecutive trials was 20, 22, or 24 seconds, as determined by chance Fifty trials were given each day and the 20th, the 29th, the 40th, and the 50th trials were controls. In these four trials each day the light stimulus was *not* followed by shock Prior to the training each rat was given five presentations of the light alone The first presentation appeared to accelerate the breathing of two of the rats Otherwise no response was observable

Results Five rats were used Each of these always responded to the brief electric stimulation by a scampering movement which produced a violent displacement of the stylus of the Marey tambour During the first day or two each made many efforts to escape from

the apparatus. But neither these nor any other observed movements were at first correlated with the presentation of the light stimulus. Three of the rats never showed such correlation. The fourth, after over 200 trials, usually lowered its head and braced its feet upon receiving the light stimulus, apparently in anticipation of the shock. In the control tests it usually raised its head and often shifted its feet a moment after the time at which the shock was usually given. The fifth rat, after 150 trials, frequently raised itself on its hind legs and tried to push its way out when the light was given but rarely did so between trials. Since other experiments, using sound instead of light, which were being conducted simultaneously, were yielding more uniform results, the light experiments were abandoned at this point. The three rats whose behavior seemed entirely unaffected by the light may have had impaired eyesight. They did not appear to be blind, however, and, since the light stimulation was of considerable intensity, it is probable that the cause for their lack of response to it must be sought elsewhere.

2. The next experiments were conducted exactly like those with light except that the buzzer replaced the 100-watt lamps. The 2.5-volt pilot lamp remained lighted throughout as before. The interval between the two stimuli was again one second. The five preliminary presentations of the sound stimulus alone before training the rats produced a more definite response than had such tests with light. It was a slight jerk of the body, varying in degree from one individual to another, and a holding of the breath for 2 or 3 seconds. It was always far less violent than the response given to the 40-volt shock later used. It was interesting to note that this response to the buzzer in every case died away, either during the five presentations alone or during the early training trials. As will be seen, the sound came to affect the animals again after many repetitions just before shock. But prior to this, some process (experimental extinction or negative adaptation) had apparently rendered the buzzer temporarily ineffective.

The behavior of the animals used in this experiment will be reported in more detail. In the accompanying tables (Figure 4) are indicated the 50 trials of each day's series. Beneath each, and opposite the day in question, is shown by a numeral what behavior, if any, was detected following the sound and before the shock was given in that trial. An underline indicates that, during the between-trial interval preceding, the rat had been making an effort to escape

from the cage There is an extra space (labelled *c*) after each control trial and in this column is noted whatever behavior the animal made at the time the shock would have been given had it not been a control trial.

The numerals used and the behavior indicated by each are as follows

1 *Jump* An escape effort.

2 *Raising on the hind feet* This is less violent and has the appearance of being exploratory

3 *Raising the head but keeping the four feet on the electrodes.* This seems to be an abbreviated form of 2

4. *Body-sway* Sometimes should be called shoulder-sway. The weight is shifted from one side to the other Usual periodicity of about 1½ second, d.v

5 *Head-sway* Usually with the head lowered and at about the same rate as the body-sway.

6 *Accelerated, or labored respiration.* Sometimes detected visually, sometimes audibly also, and sometimes recorded on the kymograph

7. *Holding of the breath, or deceleration of respiration*

8. *A cry, squeak, or "chittering,"* the animal having previously been silent; or the rate of such sounds suddenly altered (usually accelerated)

9 *Lowering of the head and bracing of the feet.*

0 *Sudden cessation of some miscellaneous activity* such as "chittering" scratching turning about or raising on hind feet

Of these forms of behavior only 1, 2, and 4 were always recorded by the kymograph Whether the others were recorded depended both upon their violence and upon whether they occurred when the animal was otherwise quiet.

Again it must be emphasized that in the tables these numerals always refer to the behavior which immediately *followed* the sound stimulus. If such behavior occurs only, or almost only, at that time we may presume that the sound stimulus was a cause Inasmuch as the only behavior which occurred at all frequently at other times (except, of course, the scampering response to the shock) consisted of escape efforts and their occurrence is always indicated by the underlining of the space for the next following trial, the reader can judge for himself whether a behavior pattern is initiated by the sound Let us take rat XD as an example During the first 50

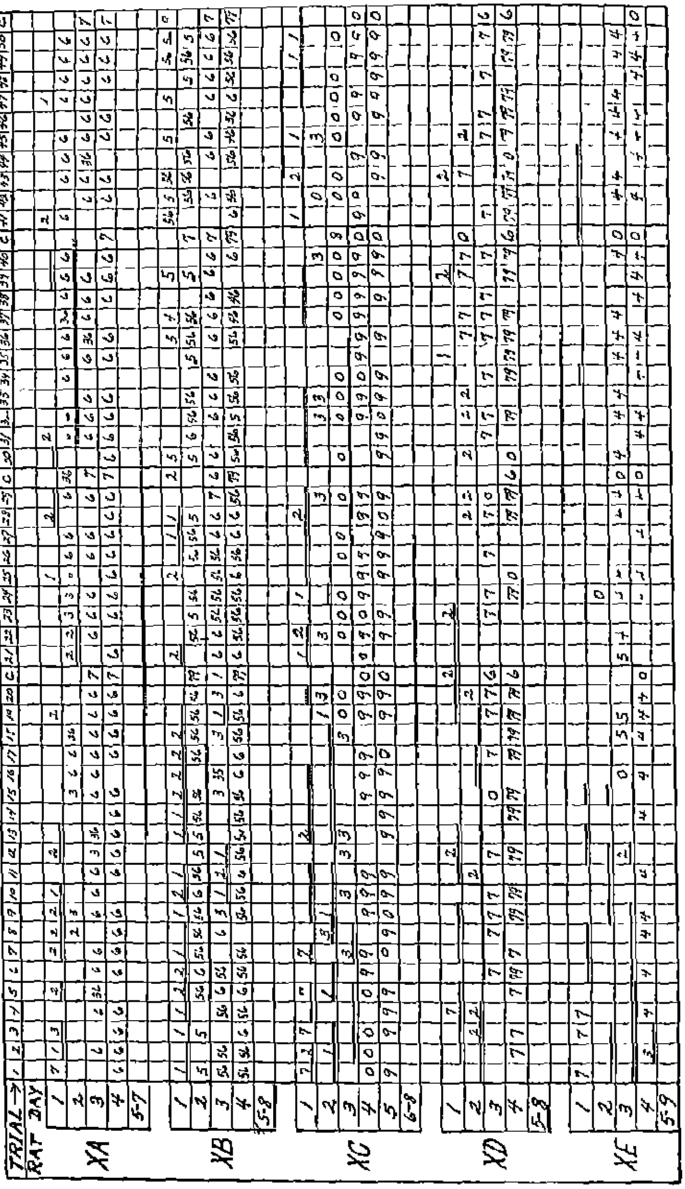


FIGURE 4
The Suspended-Cage Situation

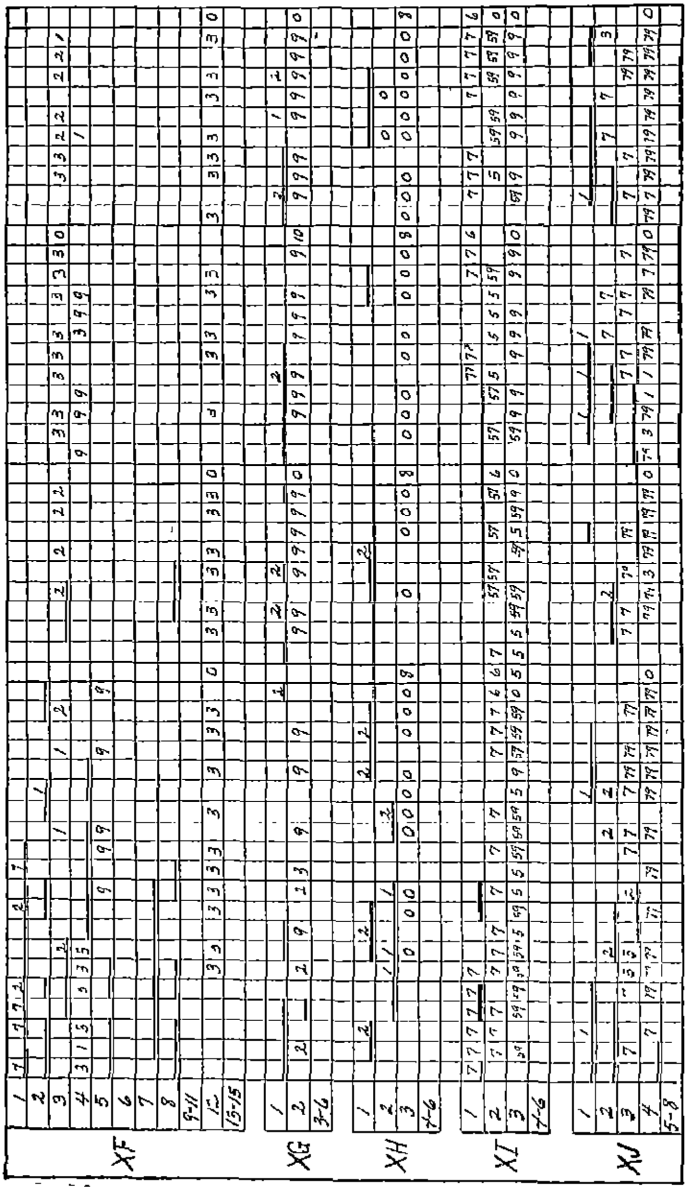


FIGURE 4 (*continued*)

trials it raised itself to its hind legs on five occasions immediately following the sound stimulus. But it is noted that on four of these occasions such escape efforts had also immediately preceded the trial. Certainly, here is no evidence that the animal responds to the sound by rising to its hind legs. But during the last 25 trials of the second day the animal does not once show such behavior between trials, while on six occasions it does so immediately after the sound is given. We suspect the sound of playing a part in initiating this behavior and, at the same time, we note other evidences that the sound is becoming a factor in the determination of the animal's behavior. On five occasions it holds its breath at the sound stimulus.

The days of training after a form of behavior had become quite stereotyped have been omitted from the table since they would be almost repetitions. For example, rat XA made records on days 5, 6, and 7 almost identical with its record on the fourth day.

Results. It is suggested that the table (Figure 4) be followed in conjunction with the verbal description.

Rat XA at first made many efforts to escape from the cage. As the table shows, it was frequently on its haunches when the buzzer sounded and remained in that position (2) until the shock. When the shock was given it regularly dropped to all fours. During each of the control tests of the first and second days it merely continued whatever activity it was engaged in. On the second day, however, the buzzer was definitely responded to, both by raising the head and by accelerated respiration. This rat practically always breathed rapidly in response to the buzzer on the third day and in three of the four control tests its breathing gradually returned to normal when the shock failed to be given after the one-second interval. On the succeeding 200 trials the animal's behavior was quite uniform: acceleration of respiration for the buzzer, followed by deceleration after the shock, or, in the control tests, after the moment had passed when the shock was usually given.

Rat XB tried persistently to escape during the first 28 trials. After this it remained constantly on the floor of the cage where its only pronounced movement was a head-sway—which was given *only* after the buzzer was sounded and ceased with the shock. Labored respiration frequently accompanied this movement. The fourth, fifth, and seventh control tests suggested that the rat "expected" a shock. In all control periods from the third day on the

animal held its breath and usually lowered its head and braced, only relaxing after the moment of the "expected" shock had passed

Rat XC required over two days' training to end its escape efforts. During these days it chirped rhythmically most of the time. Beginning with its 19th trial of the third day, it frequently became suddenly silent when the buzzer sounded (0) On the fourth day this cessation of chirping was usually accompanied by the bracing response. The first successful control occurs on the 20th trial of the fourth day. It relaxed its position and raised its head when the shock failed to be given at the usual time From this time on its behavior remained almost uniform

The escape efforts of XD consisted of raising itself on its hind feet and nosing about the lid of the cage At first this behavior occurred mostly during the between-trial interval During the second day, however, the animal usually remained quiet during this period, but, upon hearing the buzzer, it would stand and attempt to push its way out with its nose. The shock always abruptly terminated these attempts Beginning with the later part of the second day this behavior was replaced with another sort. Upon hearing the buzzer, the rat would more frequently hold its breath and remain tense until the shock was given. Occasionally this behavior would interrupt other random movements such as turning about in the cage, yawning, etc Although in the table the symbol *79* suddenly replaces the symbol *7*, actually the transition from one to the other was gradual The animal instead of merely holding its breath came to brace itself more and more tensely and with head lowered until finally the nose all but touched the electrodes at the gap. During 300 trials the behavior was not further altered to an appreciable extent.

Rat XE made less persistent efforts to escape than did the others. It also reacted less violently to the shock. During the first 25 trials of the third day the shock was gradually increased to 54 volts and it was then more effective. Almost at once the animal adopted a behavior pattern which it maintained throughout the remainder of its work. When the buzzer was sounded it shifted its weight from one side to the other in a swaying motion and when the shock was given it lifted its feet gingerly In the control tests, when the shock was omitted, it ceased swaying soon after the usual moment for the shock

Rat XF was excessively inactive. Feeble escape efforts occurred

on the first two days On the third day these efforts, particularly the act of standing on its hind legs and nosing about the top of the cage, became restricted to the period immediately following the sound. The shock did not interrupt this activity as brusquely as it did with the average rat. This behavior was apparently abbreviated to the mere lifting of the head in response to the sound But on the following day even this died out, and, except for a few half-hearted "bracings" for the shock (*9*), there is little to report until the 12th day. During this period the rat became increasingly sluggish in spite of the gradual increase in the strength of the shock used. On the 12th day the head-raising reappeared and now, when a control trial occurred, it lowered its head a moment after the time when the shock would have been given. This behavior continued with little variation during the remaining three days during which the rat was tested.

Rat XG arrived at a stereotyped form of behavior which closely resembled that exhibited by XC. It was the rat which most promptly reached a uniform type of response, changing little after the second day.

Rat XH made a periodic chirping sound from the beginning. During the third day this became the basis for a pattern which remained uniform throughout the remainder of its trials The chirping would be interrupted by the sound and resumed again soon after the animal had settled down after its scampering response to the shock. In the control trials it resumed its chirping two or three seconds after the shock was supposed to have been given

Rat XI gave evidence of attention to the sound far earlier than did the average animal. During the last 15 trials of the first day it seemed evident that the animal held its breath when, and only when, the sound was given, and resumed normal respiration only after the shock The control tests showed resumption of normal breathing after the time for the shock had come and gone. During the next two days the rat passed through transition forms of response, first swaying its head slightly while holding its breath, then lowering its head and bracing while swaying, and, finally, merely bracing with lowered head This behavior remained relatively unchanged through the remainder of the tests.

Rat XJ first showed an influence of the sound stimulation during the last part of the second day by holding its breath when the sound was given This was frequently accompanied by lowering the head

and bracing the feet, a combination which became established the following day to remain uniform thereafter

Situation III—The Parallel Bars

When it was evident that in Situation II, as in Situation I, we had failed to discover true conditioning (since none of the several responses which had eventually become linked with the secondary stimulus had been the response given to the primary stimulus), a new situation was devised. Two monelmetal rods, each about 1/16th of an inch in diameter and 12 inches long, were mounted parallel to each other and separated by 1 inch. The mountings were small slabs of bakelite, one at each end of the rods They were suspended in a horizontal position by wires from the bakelite slabs to the ceiling. To prevent rats from escaping from these rods by climbing the wires, two large pieces of thick cardboard, each with two holes cut for the rods, were mounted as shields. The pieces of cardboard extended sufficiently above, below, and to the sides of the rods to prevent the rat from reaching their edges. The surface of the cardboard was too smooth to be climbed Thus the only way a rat could escape from the horizontal bars was by dropping or jumping to whatever happened to be below. Each of the two bars was connected to one of the electrodes of the electrostimulator Immediately above the bars, and at a distance of 2 feet from them, was suspended the electric buzzer that had been used in the previous experiments.

Preliminary work (with rats previously used in the suspended-cage problem) showed that rats are loath to abandon the bars unless they can, when hanging, touch the floor beneath Sounding the buzzer failed to make them leave their perches, but for a sufficiently intense electric shock they hastily jumped. At the beginning of the experiment the bars were only 2 feet above the floor

1 Each rat was given five sound stimulations while on the bars. This caused none of them to jump or make other notable response Then rats were trained with a five-second interval between the one-half-second sound stimulus and the shock. The shock was not turned off automatically, but remained on until the animal had "solved" the problem On the first trial only six of the eleven rats used made the response to the shock that the experimenter had anticipated, i e, that of leaping off the bars. The others solved the problem, at least to their own satisfaction, in any of several ways by sitting with both hind feet on the same bar and sometimes steadying themselves with

a fore paw on the cardboard shield; by sitting on a single bar and leaning with the tail on the other (the tail appears to be insensitive); by hanging from one of the bars with either the fore or the hind paws and finally letting go When the six rats which jumped were given further training (10 trials the first day), five of them continued each time to wait for the shock before jumping, but the sixth would jump almost as soon as placed on the bars, not waiting for either sound or shock. For this animal the bars were raised, 6 inches per trial, until at a height of 4 feet it awaited the shock The same procedure was followed later when three of the other rats started jumping as soon as placed on the bars Of the six animals, one would never jump off until after the shock, although given 250 trials Another animal, after the 95th trial, would jump off the bars at whatever height the experimenter cared to place them. Naturally there was a limit to the height which could be used as it was not our desire to injure the animals. The remaining four animals, after from 90 to 140 trials, adopted the quite uniform habit of resting on the bars until the sound stimulus was given. Then they did not jump hastily off as they did for the shock but hung for a moment, usually by the hind paws, and dropped to the ground

Obviously, this was not the leaping-oft-into-space response which each of these animals always (both before and after training) made in response to electric shock The response that had been grafted onto the sound stimulus had not been snipped from the shock stimulus

Situation IV—The High Fence

The final situation to be reported here was not developed deliberately but grew out of a casual observation. In the preliminary work on the association span (situation described in the first paragraph of this article) a fence was used which presented a rather wide gap between its lower edge and the monelmetal bars which formed the floor of the cage. On several occasions animals attempted to squeeze their way beneath the fence rather than to jump it. The fence was modified and the work continued without further difficulty from this source. But afterward the fence was rebuilt so as to permit an animal to scramble beneath, if it so chose, rather than to leap over it Since the animals used were not uniform in size, the fence was so devised that it could be raised for a large rat and lowered for a smaller one The effort was made to keep as constant as possible

the ease with which each of the rats could scramble beneath the fence. This could only be accomplished approximately, of course, as the results will show. For the average rat the top of the fence was 1½ inches above the bars which formed the floor, while the gap between the floor and the bottom of the fence was 1⅛ inches.

The training given the animals of this group resembled that used with the one-second group as described previously (1) Whether an animal reached the opposite side of the fence by going over or under it mattered not a whit so far as solution of the problem was concerned. Whichever it did, it avoided the shock (if the response was made to the sound) or ended the shock (if it had waited until the shock stimulation was given). All eleven of the animals trained eventually did learn to avoid the shock for at least six consecutive trials and usually more But there was quite a difference in the methods employed Five of the animals always, or very nearly always, leaped over the fence whether in response to the shock or to the sound. It may be that the fence had not been adjusted so high for them (relative to their sizes) as it had for the other animals. Two of the animals would go over or under the fence, indifferently, and which method they selected did not seem to depend upon whether the response was made to sound or shock

The records for the remaining four animals must be given in more detail (Figure 5). The symbols in this table are as follows.

1 *Abortive or misdirected behavior.* This consists of efforts to escape from the apparatus by jumping

1*a*—indicates that after jumping the animal fell to the same side of the fence

1*b*—indicates that the jump carried it across the fence

2. *Undirected behavior,* suggesting that the animal hears the sound

2*a*—indicates that the animal squeaked or "chittered" having previously been silent, or that the rate of its sounds was suddenly altered (usually accelerated)

2*b*—indicates that the animal having previously uttered such sounds rather periodically suddenly ceased

2*c*—indicates that the animal quivered or that its respiratory rate increased

2*d*—indicates that the animal ceased quivering or held its breath.

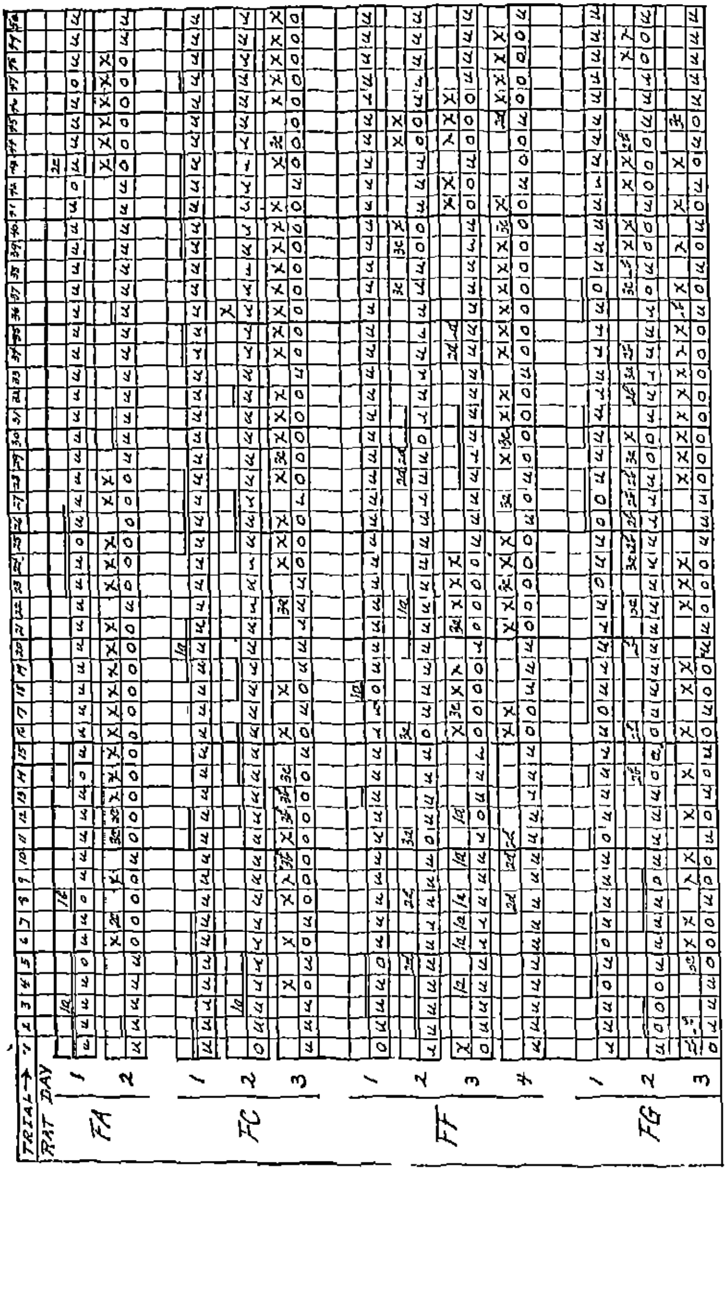


FIGURE 5

THE HIGH-FENCE SITUATION

2e—indicates that the animal ceased some other activity such as exploring or scratching.

3. *Correctly directed behavior* which is, however, *uncompleted.*

3a—indicates that the animal turned its head toward the fence

3b—indicates that it actually approached the fence and was in position to leap across.

3c—indicates that the animal made a feint at leaping across, getting at least its forefeet and head above the fence, but then hesitated and withdrew

4 *Correctly directed and completed behavior*

X—symbolizes the response of crossing the fence to the opposite half.

The underlining of a symbol indicates that escape efforts had immediately preceded. It will be noted that immediately under each of these symbols is placed either "*o*" or "*u*" The former signifies that on this trial the animal crossed over the fence, the latter that it crossed under

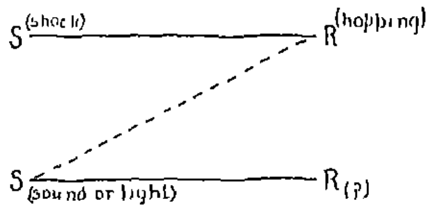
During the first 50 trials rat FA crossed *over* the fence only 6 times, 3 of these occurring at a time when it was making frequent efforts to escape from the apparatus by jumping upward On the next 50 trials the rat crossed after the sound and before the shock 22 times and on each of these occasions it jumped *over* the fence Of the 28 trials when it did not cross in response to the sound, but only when the shock was given, it went under the fence 23 times In 2 of the 5 exceptional trials it had made a feint at crossing (*3c*) in response to the sound and was poised with foot on fence when the shock was given It is evident that FA responded to the shock by scampering beneath the fence, but that when the sound was sufficient to cause it to cross it did not make the same response but quite a different one, that of crossing *over* the fence The records of rats FC, FF, and FG support the same conclusion and need not be discussed in detail. But perhaps one comment should be added The crossing beneath the fence can best be described as a hasty scamper The crossings over the fence were far more smooth They gave the impression of being deliberate, though not slow

Summary

Perhaps the most useful summary would consist of the presentation, for each of the situations devised, of a "conditioned response" diagram for that situation The contrast between these and the

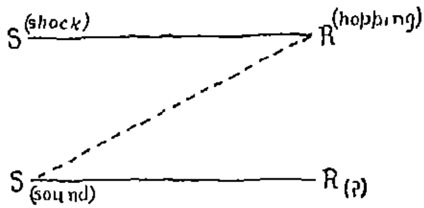
actual results will be obvious The dotted line in each case indicates the expected conditioning.

Situation I—The Box



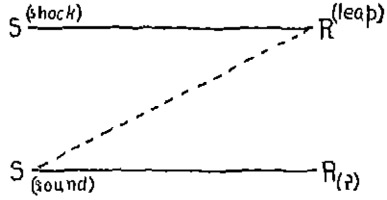
But actually no rat, even after 1000 trials, responded to the sound or light by hopping. The secondary stimuli did come to affect the animals' behavior but not in this way. The most frequent responses given to them were change of respiration rate and lowering of the head.

Situation II—The Suspended Cage



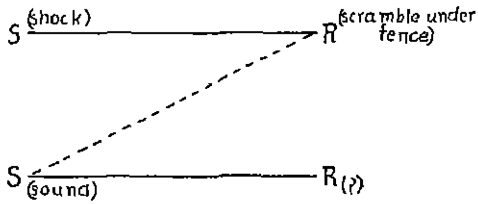
The comment given for Situation I applies also to this situation. The objective records (kymograph) of reactions to the shock and of those to the sound were utterly different, even in the case of a rat which had had 750 trials.

Situation III—Parallel Bars



Actually the rats did not learn to leap off the bars in response to the sound but to hang from one of the bars and drop downward as gently as possible.

Situation IV—The High Fence



The four rats which quite consistently scrambled under the fence in response to the shock did learn to get to the other side of the fence in response to the sound—*but by leaping over it.* Here there can be no question as to whether the difference between the two modes of crossing could have been imagined by the experimenter

Conclusions

Certain conclusions do seem warranted even though the purpose of this work was purely exploratory

1. Even though a rat be placed in a situation wherein all stimulation is relatively uniform with the exception of two potent stimuli, and even though these stimuli be oft repeated in close temporal juxtaposition, it does not necessarily follow that the response to one of these stimuli can ultimately be aroused by the presentation of the other

2 When a rat is placed in such a situation it may happen (as in Situations III and IV) that the training will result in a profound alteration in the response to one of these stimuli so that the new behavior aroused by this stimulus *produces the same end-result* as that produced by the response to the other stimulus (In the two situations mentioned, the sound came to arouse a response which removed the animal from the electrodes just as effectively as did the response which was made to the shock itself.)

3. Even though the end-result of these two responses may be the same, it does not follow that the neuromuscular mechanism employed is the same In several cases (certainly in Situations II, III, and IV) the responses are so clearly unlike that it is obvious that the response originally made to one stimulus has *not* been transfered to the other, but that an entirely new response has developed.

The conditioned response hypothesis has been just as susceptible to over-application and mis-application as have, for example, the

Freudian principles One should not conclude that the hypothesis is faulty but rather that it should be employed far more critically, and that it should be made the subject of less speculation and more experimentation.

Reference

1. Warner, L. H. The association span of the white rat *J. Genet Psychol*, 1932, 41.

Columbia University
New York City

UNE RECHERCHE EXPÉRIMENTALE DE LA "RÉPONSE CONDITIONNELLE"

(Résumé)

En faisant des expériences sur la plus grande durée de l'association chez le rat blanc, l'auteur a noté un comportement qu'il n'a pu résoudre facilement en terms de la réponse conditionnelle puisque la réponse faite au stimulus original ou physiologiquement inadéquat et la réponse jointe au stimulus (supposé) conditionnel ont *paru* être qualitativement dissemblables. Cet article décrit d'autres situations expérimentales faites dans le but de déterminer si cette différence qui a semblé se montrer a été vraie Les expériences (toutes avec les rats) lesquelles ne peuvent pas être détaillées ici, ont fait conclure à l'auteur (1) que la juxtaposition temporelle souvent répétée de deux stimuli ne produit pas nécessairement le conditionnement, (2) que la juxtaposition de deux stimuli, cependant, change souvent profondément la réponse faite à l'un des deux, (3) que ce comportement nouveau produit souvent le même résultat final que celui produit par la réponse au stimulus original, et enfin, (4) que bien que le résultat final puisse être le même, il ne suit pas de là que le mecanisme neuromusculaire employé soit le même.

Warner

EINE EXPERIMENTELLE UNTERSUCHUNG DER "BEDINGTEN REAKTION"

(Referat)

In seinen Untersuchungen uber die Assoziationen weisser Ratten, beobachtete der Verfasser ein Verhalten, das nicht ohne weiteres auf den Begriff der bedingten Reaktion zuruckgefuhrt werden kann, da die Reaktion auf die originalen und psychologisch angemessenen Reize und die auf den (angenommenen) bedingten Reiz qualitativ ungleich zu sein *schienen*. Die vorliegende Arbeit berichtet uber weitere experimentelle Situationen, die entworfen wurden, um festzustellen, ob dieser scheinbare Unterschied auch ein wirklicher ist Die Experimente (alle an Ratten), die hier nicht ausführlich dargestellt werden konnen, fuhrten den Verfasser zum Schluss, 1) dass die oft wiederholte zeitliche Nebeneinanderstellung zweiter Reize

nicht notwendigerweise eine bedingte Reaktion hervorruft, 2) dass aber die Nebeneinanderstellung zweier Reize oft die Reaktion auf den einen derselben gründlich verändert, 3) dass dieses neue Verhalten oft dasselbe Endergebnis hervorruft wie die Reaktion auf den ursprünglichen Reiz und endlich 4) dass auch wenn das Endergebnis dasselbe sein mag, daraus nicht folgt, dass der benutzte neuromuskulöse Mechanismus derselbe ist.

WARNER

SOME NEW BASES FOR INTERPRETATION OF THE IQ*

From the Iowa Child Welfare Research Station, State University of Iowa, Iowa City, Iowa

BETH L. WELLMAN

Since its establishment in 1917, the Iowa Child Welfare Research Station has been following the development of the same children over a number of years. One part of this program has been concerned with the measurement of intelligence by means of the Stanford and Kuhlmann revisions of the Binet scale. A first report on school children was made in 1922 by Baldwin and Stecher (1) who gave results based on 485 mental ages for 143 children, 36 of whom had had five examinations A second report was made by them (2, pp. 53-62) in 1924 on Stanford-Binets for 105 preschool children, 13 of whom had had five examinations

The data for this report include approximately 3000 Binets on children from 2 to 14 years of age, about 1350 being first tests and 1650 being repeated tests, from two to seven tests per child Some children have been followed over a period of nine years. The data of Baldwin and Stecher are included.

The program has been essentially to test twice a year at six-months' intervals all children attending the preschool laboratories If these children have continued in the University Elementary and High Schools, they have been tested once each year. If they have enrolled in the Iowa City public schools, they have been tested at irregular, longer intervals Children who were enrolled in the University Elementary School prior to the opening of the preschool laboratories in 1921 were tested yearly. Since then new entrants to the Elementary School have been tested once only. Of our total repeated test group approximately two-thirds started in the preschool groups and one-third in the Elementary School.

The Stanford revision was used with all children from 4 to 14 years of age, with most of the two-year group and with about half

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, December 4, 1931.

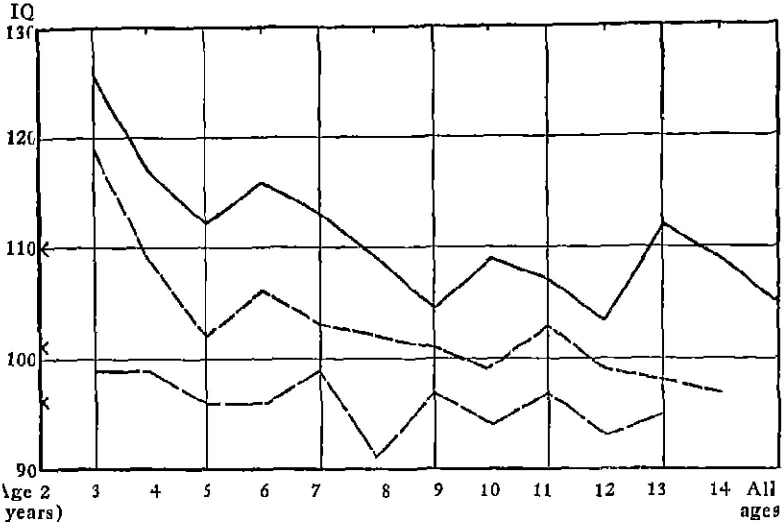


FIGURE 1

MEAN IQ ON REPEATED TESTS

1st test ————, 2nd-3rd test — — — —, 4th-7th test —— —— ——

Number of Children

Age (years)	2	3	4	5	6	7	8	9	10	11	12	13	14	All
1st test	93	172	150	176	180	95	72	69	77	67	71	50	61	1333
2nd 3rd test		112	181	162	137	107	71	53	54	60	40	30	20	1027
4th 7th test			39	98	79	35	43	52	53	42	55	46	32	574

of the three-year group, the Kuhlmann revision was used. An age group consists of children six months older and younger than the exact year which forms the midpoint. The highest age included is 14 years and 5 months

The first figure shows the mean IQ's by age on the first test, on the second and third tests combined, and on the fourth, fifth, sixth, and seventh tests combined. The curve for the first test shows a drop from the earlier to the later ages, the mean IQ at 5 years being 116 5 and at 14 years, 93 8, a difference of 22 points This curve follows much the same trend as Freeman's (3, pp. 190-191) smoothed curve based on the 905 cases used by Terman in his standardization of the test, although our curve is at a higher level Our curve is from 4 to 11 points above the Freeman-Terman curve, except at 14 years, where it is 3 points below. The curve for the second and the third tests combined follows the curve for the first test but at a still higher level The means for the fourth to seventh tests combined

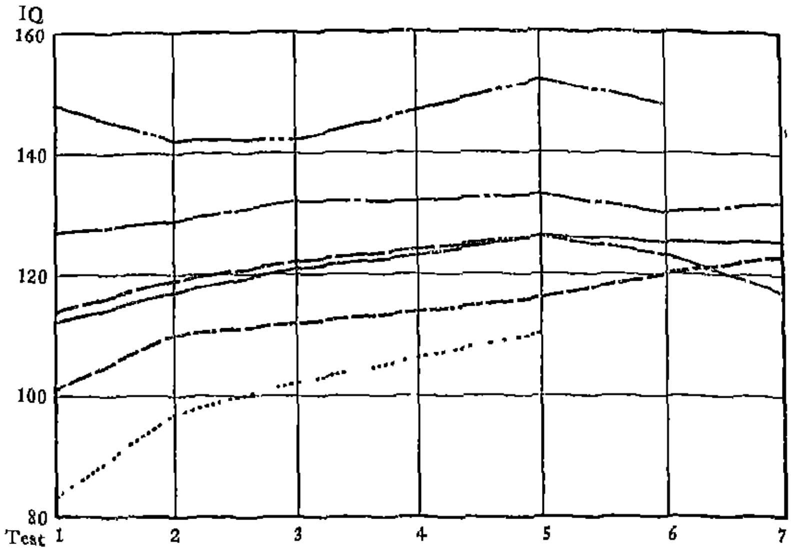


FIGURE 2

MEAN IQ ON REPEATED TESTS FOR CHILDREN CLASSIFIED INTO IQ LEVELS ON FIRST TEST

Below average (89 and below)
Average (90-109)
Superior (110-119)
Very superior (120-139)
Genius (140 and above)
All

Number of Children

Test	1	2	3	4	5	6	7
Below average	41	41	20	11	7		
Average	246	246	135	87	57	37	11
Superior	155	155	90	77	47	30	11
Very superior	148	148	88	61	43	25	11
Genius	34	34	26	16	14	6	
All	647	647	395	259	174	105	37

form more nearly a straight line, but are above the previous tests. The mean IQ on the first test for the entire group, 1333 children, was 110, for 1027 cases on the second and third tests the mean was 119; and for 574 cases on the fourth to seventh tests, 124.

The children were classified into IQ levels on the basis of their first test, using the divisions and terminology of Terman. There were 647 children who had had two tests, 395 children who had had three tests, 259 having four tests, 174 having five tests, 105 having six tests, and 37 having seven tests The whole group, shown by

the solid black line in the second figure, made gains with each succeeding test up to the fifth test, in spite of the fact that the children were growing older and a decrease should be expected if the drop in the curve on the first test shown in the previous graph is valid and representative. The mean intervals between first and second tests was 12 months, between second and third tests 12 months, and between succeeding tests 9 months, the mean interval between the first and seventh tests was 5 years and 2 months. There were children with 9-years' interval between their first and fifth, sixth, or seventh tests, and others with intervals of 7½ years between their first and second, third, or fourth tests. The mean age at first test was 53 to 63 months for the various groups. No child who had seven tests was more than 9 years at the time of the first test, but the two test groups included some 14-year-olds on the first test.

The greatest gains were made by the group starting below average (shown by the dotted line at the bottom of the second graph), who had gained 28 points by their fifth test and had passed into the superior classification. The next greatest gain was by the average group, who had gained 22 points from their first to seventh test and had passed into very superior territory. The superior group gained 12 points up to their fifth test and then dropped on their seventh test to 3 points above their starting-point. The very superior group came out 5 points ahead of their starting-point and the genius group ended just where they began.

There were significant gains from first to second test, that is, the ratios of the differences in means to the standard deviations of the differences were 3.0 or above for all classifications except the very superior and genius. For the very superior group the ratio is 2.4, giving a probability of 99.2 chances out of 100 that there was a real gain greater than zero from the first to second test. For the whole group there was a significant gain from second to third tests and also from third to fifth tests.

The superior group stayed significantly higher than the average group until the sixth test, where the ratio drops to .9. The very superior group stayed significantly higher than the superior group up to the sixth test, where the ratio drops slightly, to 2.5 (but still with 99.4 chances in 100 that there was a true difference). The genius group was significantly above the very superior group throughout. Thus, though the children starting out in one level are likely to pass into higher classifications on later tests, those who started out at the highest levels remain highest.

The third figure gives the limits of one standard deviation above and below the mean for the different classifications and also indicates the range of scores The standard deviations tend to increase with succeeding tests up to the fourth test. For the below-average and average groups the lower limit of the standard deviation is lifted with succeeding tests. By the time the average group has had seven tests, at least two-thirds of the group are classified as superior and very superior. By the second test some of the original average group are breaking over into the genius group. Two children starting not more than 19 points apart have become by the fourth test 80 points apart. Of the group starting as geniuses two-thirds are within the limits of 130 and 153 on their second test, although one child has dropped to 113.

If the drop in IQ on first test with age is due to faulty standardization of the test rather than to selection of cases, and children keep their same relative levels of ability, decreases in IQ on later tests rather than increases should be expected. In order to correct for the

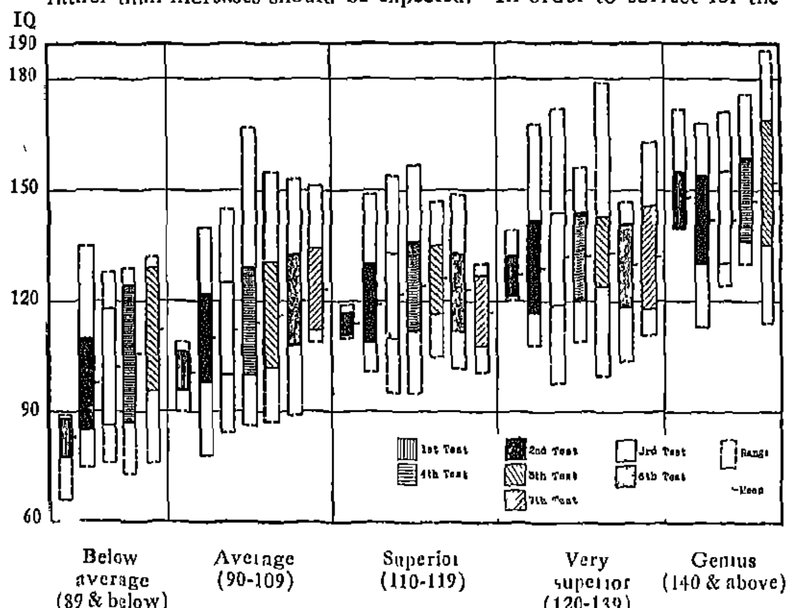


FIGURE 3

S D. Limits and Range of IQ for Children Classified into IQ Levels on First Test

unequal mean IQ's at the different ages, percentiles of IQ's were computed on the first test for each age, and each IQ on a later test was transmuted into its corresponding percentile as of first test for that age. If our children at the different ages represent similar samplings of the population, this method makes all tests and ages directly comparable. The curves shown in Figure 4 are quite similar to the IQ curves The whole group has risen from the 48th percentile on the first test to the 82nd on the seventh test. The below-average group has changed from 5th to 56th percentile, the average group from 25th to 74th, the superior group from 54th to 76th, the

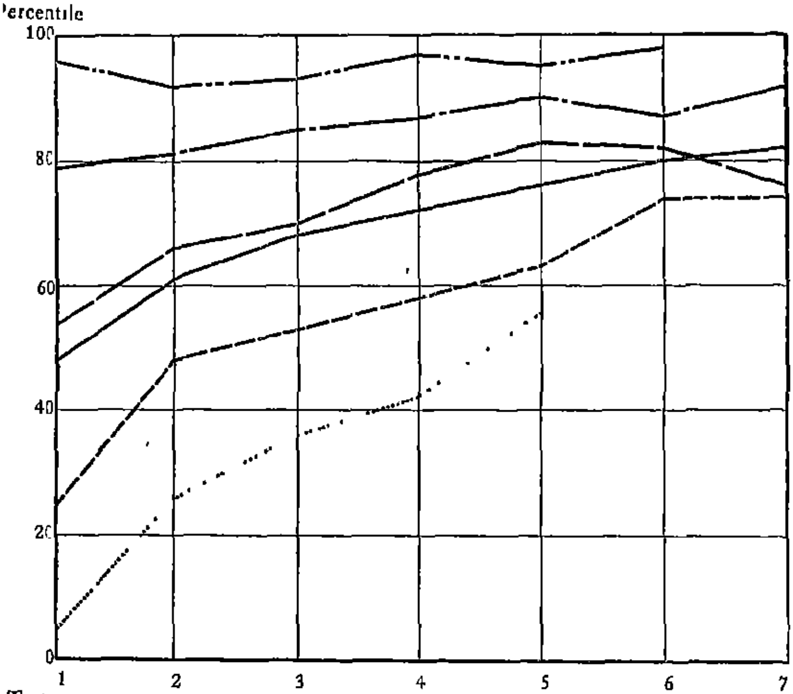


FIGURE 4

PERCENTILE VALUES (AS OF FIRST TEST) ON REPEATED TESTS FOR CHILDREN CLASSIFIED INTO IQ LEVELS ON FIRST TEST

Below average (89 and below)
Average (90-109) — — — — —
Superior (110-119) —— —— ——
Very superior (120-139) —— —
Genius (140 and above) —— — —
All ——————

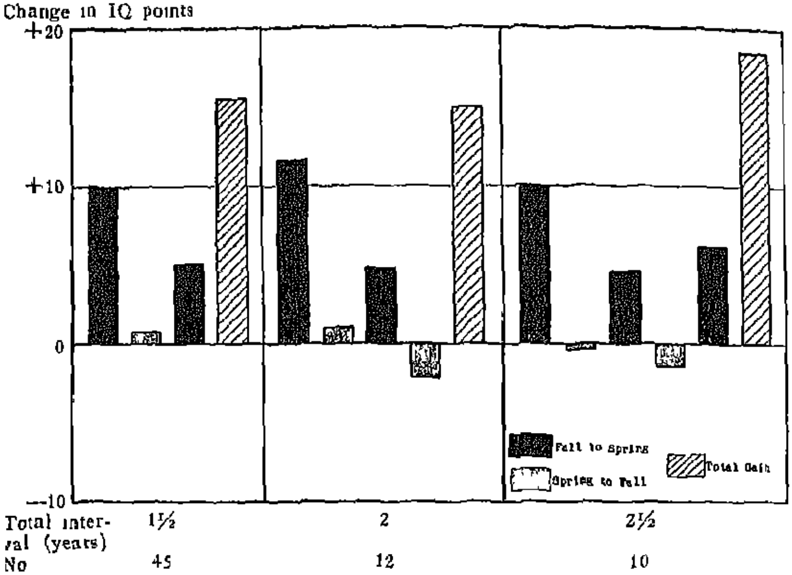


FIGURE 5
IQ AND PRESCHOOL ATTENDANCE

very superior from 79th to 92nd, and the genius group from 96th to 98th.

The next two figures present some partial evidence as to what it is that is responsible for these gains. As has been mentioned before, our preschool children receive two tests a year, one in the fall and one in the spring so long as they remain in the preschool groups. The bars on Figure 5 show the gains in IQ points made from fall to spring and from spring to fall for the same children. On the left are shown the results for 45 children, two, three, and four years of age on their first test in the fall, given either the week before they entered the preschool laboratories or within two months after entrance. The mean interval between testing was 6 months, the range being from three to eight months, with 86% falling at between five and eight months. These children attended preschool groups from fall to spring, did not attend from spring to fall, but did attend the following year and were given tests in fall and spring. The spring tests were given mostly in April and May, although a few fell in March. This group of 45 children gained 9.9 IQ points from the first fall to the first spring, less than a point from that spring to the next fall, and gained 5.0 points from the second fall to the second spring. Their total gain over one and one-half years, shown by the striped bar, was 15.6 points.

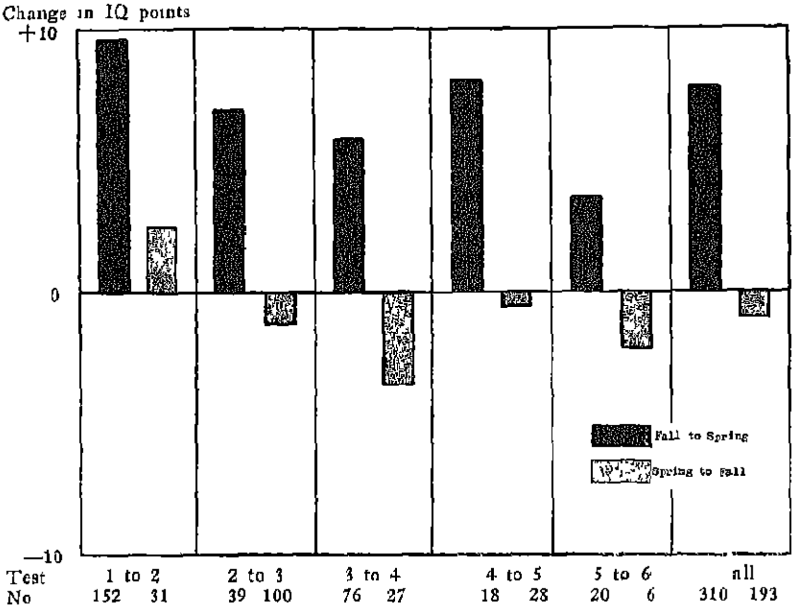


FIGURE 6

IQ AND PRESCHOOL ATTENDANCE

They began with a mean IQ of 112 8 and ended with a mean of 128 4

The second group of bars shows the gains for 12 of the 45 children who were also tested the following fall Their second spring to fall tests showed a loss of 2 2 points The third group of bars shows the gains for 10 of these 12 children who were tested also the third spring Their scores show gains from each fall to spring and failure to gain from spring to fall Their total gain over two and one-half years was 18 6 IQ points

Six explanations for these differences in gains may be advanced (1) That the differences are true seasonal differences This theory presumes a closer relationship between bodily condition and mental functioning than has hitherto been uncovered Furthermore, it does not explain how the aggregate gains came about (2) That the differences in gains are more apparent than real However, the ratios of the differences to the standard deviations of the differences indicate that there is practical certainty of a real difference between the means on the first fall and first spring for the 45 children, and between the means on the first and last tests for all three groups (3) That

Binet tests of preschool children are unreliable, and that these differences can be accounted for on this basis The correlations of IQ's between fall and spring tests for the 45 children are .84±.03 and .85±.03, while between spring and fall the correlation is .73±.04. (4) That the spring IQ's are inflated because of indirect coaching effects of experiences through which the children pass while in the preschool laboratories, which are lacking when the children are at home. (5) That the later IQ's are inflated due to practice effects from the taking of the tests. The fall to spring gains represent changes from first to second tests, third to fourth, and fifth to sixth, while the spring to fall tests represent second to third and fourth to fifth Possibly, a spurt due to practice is followed by a plateau, which happens to come between spring and fall. (6) That these gains are associated directly with preschool attendance or non-attendance.

In the next figure are compared the changes from fall to spring and spring to fall when the number of tests is kept constant. The children in the spring to fall group here may or may not have attended a preschool group for the six-weeks summer session The fall to spring group has gained even between their fifth and sixth tests, while the spring to fall groups have failed to gain except for 2.5 points from their first to second test. If these gains had been due strictly to practice effects from the mere taking of the tests, they should have been equal for the fall to spring and spring to fall groups between any given tests. The two bars at the extreme right represent changes over six-months' periods irrespective of how many previous tests had been taken. There is a gain of 7.8 points from fall to spring for 310 cases and a loss of .9 of a point from spring to fall for 193 cases.

The explanation that seems best to fit the facts is that preschool attendance, at least in the laboratories of the Iowa Child Welfare Research Station, causes a rise in IQ, the rise being cumulative from year to year, and sustained throughout the school years when the children are in the environment provided by our University Elementary School. When the same children are home over comparable intervals they fail to gain, although maintaining their higher level.

The question still remains whether these increased IQ's are real or inflated Can these children accomplish or perform other mental tasks in accordance with what would ordinarily be expected from children of their IQ levels? Do our children who were in the

beginning average group five years later act mentally as children do who are very superior? Or is the increase in IQ only, and not in intellectual ability? If the IQ is unduly inflated, corrections should be made for later tests whenever they are used Our data furnish a basis by which such a correction may be made for our group of children.

References

1 Baldwin, B T, & Stecher, L. I Mental growth curve of normal and superior children *Univ Iowa Stud : Stud Child Welfare,* 1922, 2, No 1. Pp 59

2 ——— The psychology of the preschool child. New York Appleton, 1924. Pp 305

3 Freeman, F N, Holzinger, K. J., & Mitchell, B C The influence of environment on the intelligence, school achievement, and conduct of foster children *27th Yrbk Nat Soc Stud, Educ*, 1928, Pt 1, 103-217

State University of Iowa
Iowa City, Iowa

QUELQUES BASES NOUVELLES POUR L'INTERPRETATION DU QI

(Résumé)

Un grand groupe d'enfants testés très souvent de nouveau pendant une période d'années a montré des accroissements marqués des QI On a employé la révision Stanford du test Binet, supplémentée de la révision Kuhlmann aux âges moins avancés.

Le QI moyen du premier test pour 1333 enfants âgés de deux à quatorze ans a été de 110, celui du deuxième test et du troisième pour 1027 enfants a été de 119, et celui des tests quatre à sept a été de 124 Quand on a groupé les enfants selon le premier test, on a trouvé que dans une période de cinq ans le plus grand accroissement, 28 points du QI, a été fait par le groupe au-dessous le groupe moyen, et le deuxieme plus grand accroissement, 22 points, a été fait par le groupe moyen Le groupe supérieur a gagné 12 points et puis est tombe à 3 points, le groupe très supérieur a gagné 5 points et le groupe des "génies" est resté constant

L'assistance des enfants d'âge prescolaire aux laboratories de la Iowa Child Welfare Research Station semble causer un accroissement du QI, l'accroissement étant cumulatif d'annee en annee et soutenu pendant les années scolaires quand les enfants assistent à l'Ecole elémentaire de l'Université d'Iowa Puisque les enfants qui avaient subi le même nombre de tests ont gagné pendant les périodes d'assistance préscolaire et n'ont pas gagné chez eux, les accroissements ne peuvent pas être attribués aux effets de l'exercice

Wellman

NEUE GRUNDLAGEN ZUR DEUTUNG DES INTELLIGENZ-QUOTIENTS (IQ)

(Referat)

Bei einer grossen Gruppe von Kindern, die im Laufe mehrerer Jahre mehrere Male wiedergeprüft (retested) worden waren, zeigten sich ausgeprägte Erhöherungen der Intelligenzquotienten. Es wurde die Stanford'sche Revision der Binet Tests, unter Hinzufügung der Kuhlmann'schen Revision in den Jüngeren Altersgruppen, verwendet.

Der durchschnittliche Intelligenzquotient war, für die erste Prüfung, an 1333 Kindern zwei bis vierzehn Jahre alt, 110, an 1027 Kindern für die zweite und dritte Prüfung 119, und an 574 Kindern für die vierte bis siebente Prüfung 124. Wurden die Kinder dem ersten Versuch entsprechend klassifiziert so zeigte es sich, dass im Verlauf eines fünf-Jahr-langen Zeitraums der grösste Gewinn (gain)—28 Pointe—von der unter-durchschnittlichen (below-average) Gruppe erzielt wurde. Letzterer folgte die mittelmässige (average) Gruppe, die 22 Pointe gewann. Die überlegene (superior) Gruppe verbesserte sich mit 12 Pointe, worauf die Verbesserung bis auf 3 fiel. Die hoch überlegene (very superior) Gruppe verbesserte sich mit 5 Pointe, und die Durchschnittsresultate der Gruppe der Genies blieben unverändert.

Der Schulbesuch in der Pflegeschule und im Kindergarten (preschool attendance) der Laboratorien des Iowa Child Welfare Research Station verursacht augenscheinlich eine Erhöherung des Intelligenzquotients. Diese Erhöherung häuft sich von Jahr zu Jahr an, und erhält sich durch die Jahre hindurch in denen die Kinder in der unter Leitung der Staatsuniversität Iowa stehenden Elementarschule eingeschrieben sind. Da Kinder, die eine gleiche Zahl von Prüfungen durchgemacht hatten, innerhalb der Zeiträume während der sie die Vorschule besuchten, Besserungen erzielten, und zu Hause keine erzielten, kann man die Besserungen den Einwirkungen der Übung nicht zuschreiben.

WELLMAN

ISOLATED ACTION OF COMPOUND STIMULI IN A LOCOMOTOR HABIT OF RATS*

From the Psychological Laboratories of the University of Virginia

WAYNE DENNIS AND J. M. PORTER, JR

INTRODUCTION

We are interested in a locomotor habit with compound stimuli because of its bearing on the sensory control of the maze habit After being trained under normal conditions, rats have successfully run the maze when blind, when anosmic, and when deaf (5, 7) Only seldom have all of these deprivations occurred in the same animal, but the success of the animals under these exteroceptive disabilities has led in some cases to the conclusion that the maze habit is normally controlled by touch and the proprioceptive impulses (5, 7). The success of the rat after section of some of the afferent kinaesthetic tracts has led in another case to the conclusion that maze behavior is controlled by a non-sensory intraneural mechanism (3). This theory cannot be directly tested because with all receptors rendered inoperative the animal cannot locomote But in opposition to each of the theories mentioned above we urge the possibility that maze behavior may normally be controlled by several receptors, any one of which may be effective in the absence of the others The experimenters who have used the method of deprivation have shown that certain stimuli are not essential, but they have not shown that they are ineffective when the animal is deprived of all other directive stimuli.

As long ago as 1903 Pavlov showed that either the sight or the smell or the sound of food alone excites a conditioned salivation in dogs when these have been present as compound stimuli throughout the training period (4, pp. 47-60) We should also mention the study of Bowditch and Southard (1) upon various sensory controls of human manual aiming, and the work upon sight and smell in the finding of flowers by bees (reviewed by Washburn, 6) Wylie's study (8) has some similarities to these These studies demonstrate that normal compound operation of the senses, or portions of the environ-

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, November 2, 1931.

ment, and fairly accurate isolated operation are not incompatible. It seems worthwhile to determine whether the same situation exists in rats in a learned locomotor response similar to the maze habit. The senior author has reported briefly an experiment of this nature (2); the present article is a more detailed account of a repetition of this preliminary experiment which confirms the earlier results. We shall attempt to show that stimulus objects which have been compounded during training are effective in isolation, and we shall present some evidence to show that compounded receptors are effective when activated singly.

Apparatus

The maze habit itself cannot be used to demonstrate the isolated operation of compound stimuli because in the maze the possibility that touch and proprioception are the controlling impulses cannot be removed. We must therefore use a response which we know to be controlled in its directional aspects by the exteroceptors. This requires a discrimination apparatus.

We desired a discrimination apparatus which would present a large negative area instead of the single negative compartment of the Yerkes-Watson box. This negative area was needed because we planned to present in contradictory directions two stimuli which had previously led in the same direction in order to test their comparative strength. If this were done in a box which presented only two alternatives, equal choice of the two alternatives would not reveal whether the effectiveness of the stimuli was real but equal or whether the stimuli were completely ineffective. By offering more opportunities of choice, this dilemma was avoided.

The apparatus devised to meet these requirements consisted of a circular linoleum-covered plane, 6 feet in diameter, mounted on supports which raised it about 30 inches above the floor. The rat was placed in the center of this circle and the food at some point on the edge of the circle. An iron strip, 1½ inches wide and slightly over 3 feet long, was fastened to the center of the surface by a screw which permitted the strip to rotate over the entire surface of the apparatus. This strip furnished one of the stimulus objects during the experiment. On a stool, about 3½ inches less in height than the circular surface, was placed the food box. Fastened to this stool and projecting above the circular surface was a square of cardboard, 10 inches to a side, and painted white. This served as a second stimulus object.

The third stimulus consisted in an electric buzzer placed under the seat of the stool

It will be seen that when the rat was placed in the center of the circle the experimenter was able to present the iron strip, the buzzer and the white square singly or in combination as directive cues to the food box Rotation of the food box about the circle prevented all fixed objects from serving as cues, and control tests showed that the stool itself gave no cues to its position

The apparatus was set up in a room approximately 15 feet by 25 by 15, which had no windows and which was painted a dull black All illumination came from an indirect lighting fixture placed above the center of the circle A burlap screen separated the experimenter and the rats' transportation cages from the apparatus

Animals

The experiment was begun with 27 untrained white rats of Wistar stock, approximately 90 days old After six days, 2 pregnant females and 6 rats which were very slow learners were discarded. The remaining 19 were used throughout the experiment unless otherwise stated.

Procedure

The experimental periods were 12 hours apart, training beginning at about 8 A M and 8 P M The first 12 of these experimental periods were simply periods of exploration During these each rat was placed on the apparatus individually and allowed to remain for 5 minutes Food was present in the food box, the directive stimuli were in place, and were rotated before each morning and evening period of exploration Since at the end of these 12 sessions the rats were going to the food box fairly quickly, another procedure was begun This consisted in giving each rat 5 trials during each period of experimentation Five positions 72 degrees apart were marked on the circumference of the linoleum circle and the food box and the directive stimuli were rotated from position to position counterclockwise

All the rats were run with the food box in one position, the food box and stimuli were rotated, and another run was given, etc In order to prevent the rats from always beginning their runs with the food box in the same position, no change was made between the last run of one experimental period and the first run of the next period

Guidance of the rats by the experimenter was avoided by carefully placing all rats in the center of the circle facing a fixed direction regardless of the position of the food box This direction lay midway between two of the food positions and directly away from the experimenter. As soon as each rat was placed, the experimenter stepped behind the screen and observed through a peephole.

A correct run was defined as one which did not deviate appreciably from the true pathway (not more than 2"). There was very little difficulty in applying this criterion.

If a rat ran correctly, it was permitted a bite or two of food before being returned to the cage; if it ran incorrectly, it was picked up soon after it reached the edge of the circle.

When 5 trials per experimental period were first given, 41% of the runs were correct. On the 18th and 19th experimental periods thereafter, 98% of the runs were correct At this time the tests in which we are interested were begun The special procedure in these tests can best be presented in connection with their results. The tests were given in the evening only and normal conditions were maintained in the morning periods in order to obtain a comparison record

Results

Effectiveness of Stimuli in Isolation. Each stimulus was presented singly during one experimental period. Since there were three stimuli, the tests occupied as many evenings In order to equalize as far as possible the practice effects of the tests, one-third of the rats were tested in the order of square, trail, buzzer; one-third in the order of trail, buzzer, and square; and one-third in the order of buzzer, square, and trail.

We shall present evidence later to show that the buzzer was incapable of directing the rats within the limits of the present experiment. Consequently, we shall concern ourselves here with the results of two stimuli only.

When the square alone was presented, 89% of the runs were correct. With the trail alone, 52% were correct In the morning runs of these days, 99% of the runs were correct The two separate stimuli were each effective when presented singly, the square was more effective than the iron strip, and neither was as effective as the two combined.

Opposition of the Stimuli. Following the test just described, the three possible pairs of stimulus objects were presented in opposition;

that is, the two stimuli led to different food boxes. The two food positions farthest removed from the experimenter were used and the rat faced midway between them as usual. The right-left relation of the stimuli was reversed after each run. Practice effects were equalized as in the preceding test.

When the trail and square were opposed, the rats went to the square in 96% of the trials, followed the trail in 4%, and made no errors. During the morning comparison runs, errors constituted 4% of the runs.

Control Test for Guidance from Handling. Although precautions were taken to guard against unintentional manual guidance of the rats, it seemed advisable to test for the presence of such control. Just after the rats had chosen the square 96% of the time in those runs in which the square and the iron strip were opposed, this test was repeated with each rat placed facing in the direction of the strip. In spite of this the rats went to the square in 97% of their runs. This test, together with the inability of the rats to run correctly to the buzzer or in the total absence of stimuli as shown in later tests, shows conclusively that the rats were not guided by the experimenter.

Effect of Additional Training with Trail Alone. In order to determine whether the trail alone was capable of commanding as high a degree of accuracy as the square alone, training toward following the trail was given both morning and evening for five days. Upon the last day of this special training, the rats made 93% correct runs upon the trail alone, a record which compares very favorably with the earlier 96% upon the square alone.

After some intervening training upon the buzzer taken singly (to be reported later in this paper), the trail and the square were again tested singly and were found to elicit 86% and 88% perfect runs respectively, half the rats being tested first with the trail and second with the square and half in the reverse order. These lower percentages were no doubt due to the absence of these stimuli during training upon the buzzer.

We wished next to determine whether this equality as single stimuli meant equality as opposed stimuli. The answer was found in the negative, for, in the opposition test, 81% of the rats chose the square, 19% the trail; there were no errors. Although the stimulus values proved to be unequal in this test, it should be noticed that the inequality was not as great as in the earlier test.

Isolation of Receptors. Thus far we have separated only the stimu-

lus objects We wished finally to experiment with the receptors which were stimulated by these objects It will be remembered that *the rats had found the square when it was presented in isolation* in 96% of the trials. Six of the rats, chosen at random, were deprived of vision by enucleation of the eyeballs and given 10 trials with only the square present. Only 9% of these runs were correct. A higher percentage than this should have resulted if the rats could have distinguished only the five food positions, without being able to distinguish the correct one Hence we must conclude that the square stimulated only the visual receptors On the other hand, in 10 further runs with the trail present in isolation these blind rats were 89% accurate. This shows that the trail had a large, although not necessarily an exclusive, non-visual effectiveness While it is true that for other purposes the rats had been given training upon the trail taken singly, during this training non-visual features of the iron strip were not altered in any way. We have every reason to believe that the trail would have been effective in some degree, though not as strongly, had the rats been blinded immediately after the original training period In summary, the visual and non-visual components of the compound stimuli were each effective when presented singly. Either kind of receptor could control the response when operative in isolation.

The Ineffectiveness of the Buzzer During the tests in which the stimuli were presented singly and in opposition, the buzzer attracted no larger percentages of responses than the erroneous positions about the circle After special training upon the trail alone, special training upon the buzzer alone was given for 3 experimental periods, or 15 runs. The average correctness within this time was 21%, not reliably different from the 20% one would expect by chance if the rats could select the five "fixed" points. The five rats with the best records were given 35 further trials, the remaining rats were merely fed on the food stool with all stimuli in place Of these 35 trials, 18% were correct

After all experiments which have been described heretofore, two of the blinded rats were given 700 trials with the buzzer alone From 10 to 50 runs were given in a single daily experimental period, and the training covered 35 days. The pathway was rotated as usual

The blinded rats lost even the ability of making 20% accuracy, which shows that the recognition of the food box positions had been visual. The number of correct runs, and the number of runs in

which the rats reached the edge at a point within the correct half of the circle were recorded. Within the last 80 of the 700 trials one rat was making no better than chance accuracy (41 runs in 80 within the correct half of the circle, 9 of these correct). The other rat, however, was apparently guided to some extent by the buzzer, for 59 of the 80 runs were within the proper semicircle, and 24 of these were correct. After trial 700, 80 control runs were given in which the usual procedure was followed except that the buzzer was turned off. In these runs, this rat ran 63 times in the correct half and 22 of these were correct. It is not necessary to assume, therefore, that the buzzer had been controlling this rat. The behavior of the first rat was likewise unaffected during these control tests. Apparently the rats were not deaf to the buzzer for they became active in their cages when it was turned on, but it seems certain that the buzzer had acquired no power of controlling directional behavior. This may have been because the circular plane lay between the animal and the buzzer. When this plane was removed and a human subject (blindfolded) placed his head in the former position of the rats, he had no difficulty in naming the position occupied by the buzzer. In this test each experimenter gave the other 20 trials and the blindfolded subject was led from the room while the position of the buzzer was irregularly changed. The subject was permitted to turn his head in order to locate the sound but not to incline his body.

We did not attempt to determine the stimuli which were controlling the partially successful behavior of the blind rat mentioned above, for his success developed only during the special training and hence did not affect the earlier results.

We conclude that the buzzer acquired no isolated effectiveness during the original training not because other stimuli were present, because with the amount of training there given, the buzzer was incapable of acquiring directional effectiveness.

Summary

Nineteen rats were trained in a discrimination apparatus with three stimuli present. One of these, however, exerted no control over the behavior here studied. When each of the two effective stimuli were presented singly, each was found to be effective, although neither was as effective as the two combined. The effects of the two were unequal. This inequality was even more apparent when the stimuli were placed in opposition. By further experiments it was found that

the response could be elicited both visually and non-visually, showing isolated effectiveness of receptors as well as of stimulus objects when compound receptors and stimuli were present during training

Interpretation

This experiment corroborates in the field of learned locomotor responses of rats the isolated action of compound stimuli and receptors which already has been demonstrated in other fields and with other subjects. This finding means that the continuance of a response after the deprivation of a stimulus or a receptor does not prove that the deprived stimulus or receptor has been ineffective. As a corollary to this, Watson's proof of the kinaesthetic theory and Lashley's proof of the neural-mechanism theory of the maze habit are seen to be inadequate.

References

1. Bowditch, H P, & Southard, W F. The control of aiming by sight and touch *J Physiol*, 1881, 3, 232-245
2. Dennis, W The sensory control of the white rat's response in the maze *Clark Univ Thesis Abstracts*, 1930, 1, 33-35.
3. Lashley, K. S., & Ball, J Spinal conduction and kinaesthetic sensitivity in the maze habit *J Comp Psychol*, 1929, 9, 70-106
4. Pavlov, I. P Lectures on conditioned reflexes New York: International Publishers, 1928. Pp 414.
5. Small, W S An experimental study of the mental processes of the rat *Amer. J Psychol.*, 1901, 12, 206-240
6. Washburn, M F The animal mind New York. Macmillan, 1926 Pp. xiii+431.
7. Watson, J B. Kinaesthetic and organic sensations *Psychol Rev, Monog. Suppl.*, 1908, 8, No 33. Pp vi+100
8. Wylie, H H. An experimental study of transfer of response in the white rat. *Behav. Monog*, 1919, 3, No 5. Pp. 65.

University of Virginia
University, Virginia

L'ACTION ISOLÉE DES STIMULI COMPOSÉS DANS UNE HABITUDE LOCOMOTRICE DES RATS

(Résumé)

On a entraîné dix-neuf rats dans un appareil de discrimination avec trois stimuli présents. L'un de ceux-ci, cependant, n'a pas été capable de contrôler le comportement de direction ici étudié. Quand chacun des deux stimuli efficaces a été présenté seul, chacun s'est montré efficace, quoique ni l'un ni l'autre n'ont été si efficaces que les deux combinés. Les effets des deux ont été inégaux. Cette inégalité a été même plus prononcée quand on a opposé les stimuli l'un à l'autre. Au moyen d'autres expériences on a constaté qu'on a pu faire venir la réponse visuellement et non visuellement, ce qui montre l'efficacité isolée des récepteurs aussi bien que celle des objets de stimulus quand des récepteurs composés et des stimuli composés ont été présents pendant l'entraînement.

Cette expérience corrobore dans le domaine des réponses locomotrices apprises chez les rats l'action isolée des stimuli et des récepteurs composés laquelle a été déjà montrée dans d'autres domaines et avec d'autres sujets. Cette constatation signifie que la continuation d'une réponse après la perte d'un stimulus ou d'un récepteur ne prouve pas que le stimulus ou récepteur perdu n'a pas été efficace. Comme corollaire de ceci, la preuve de la théorie kinesthésique par Watson et celle de la théorie du mécanisme nerveux de l'habitude du labyrinthe par Lashley se montrent inadéquates.

DENNIS ET PORTER

DIE ISOLIERTE EINWIRKUNG ZUSAMMENGESETZTER REIZE BEI EINER BEWEGUNGSGEWOHNHEIT DER RATTE

(Referat)

Es wurden 19 Ratten eingeübt in einem Unterscheidungsapparat (discrimination apparatus) in der Gegenwart von drei Reizen. Einer dieser Reize war aber zur Beherrschung der hier untersuchten Richtungstätigkeit (directional behavior) unfähig. Wenn jeder der zwei wirksamen Reize einzeln dargeboten wurde, ziegte sich jeder als wirksam, obwohl weder der eine noch der andere so wirksam war, wie die zwei in Verbindung. Die Einwirkungen der zwei waren ungleich. Diese Ungleichheit war noch stärker ausgeprägt, wenn die Reize einander entgegengesetzt wurden. Weitere Versuche erwiesen, dass die Reaktion sowohl visuel wie nichtvisuel hervorgerufen werden konnte. Hierdurch wurde die isolierte Wirksamkeit der Rezeptoren wie auch der Reizgegenstände, wenn zusammengesetzte Rezeptoren (compound receptors) während der Einübung gegenwärtig waren, erwiesen.

Dieser Versuch bestätigt im Bereiche der erlernten Bewegungsreaktionen, bei Ratten, die isolierte Einwirkung zusammengesetzter Reize und Rezeptoren, die schon in anderen Bereichen und an anderem Versuchsmaterial erwiesen worden war. Dieser Befund weist darauf hin, dass die Fortsetzung einer Reaktion nach Entziehung eines Reizes oder eines Rezeptors nicht beweist, dass der entzogene Reiz oder Rezeptor wirkungslos gewesen sei. Es folgt hierauf als Korrelat, dass Watson's Beweis der kinesthetischen Theorie und Lashley's Beweis der Theorie eines neurologischen Getriebes (neural mechanism theory) zur Erklärung der Labyrinthgewohnheit (maze habit) ungenügend sind.

DENNIS UND PORTER

THE EFFECT OF MOTIVATION UPON INTELLIGENCE TEST SCORES*

From Teachers College, Columbia University

JULIUS B. MALLER AND JOSEPH ZUBIN

In comparing the intelligence test scores of different individuals, or of the same individual on different tests, the degree of motivation under which an individual worked during each testing period is seldom considered. Thorndike has pointed out, among other limitations of present-day intelligence tests, that ". . . all our measurements assume that the individual in question tries as hard as he can to make as high a score as possible. . . . In general practice, however, we rarely know the relation of any person's effort to his possible maximum effort" (4, p. 228). Does the degree of motivation that obtains during the testing process affect the results? Would the score on an intelligence test obtained under normal test conditions increase if the degree of motivation were increased?

Hurlock (1) has presented data showing that motivation will affect the score appreciably. Three groups were equated for IQ on Scale B, Form 1, of the National Intelligence Test. They were then retested with Form 2 of the same test under the following conditions of motivation. One group was reproved for poor performance, the second was praised for good performance, while the third was retested without any additional motivation. She found that praise and reproof were each more effective in raising the score than mere repetition of the test.

This present study concerns itself similarly with the effect of additional motivation upon intelligence test score.[1] It differs, however, in two respects. First, it made use of a different, and probably stronger, incentive—rivalry. Secondly, it utilized the same form of the test for the initial and final testing.

The incentive of rivalry was chosen because it is present to some extent in the usual testing situation. The rivalry introduced in the

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, February 8, 1932.

[1]The authors wish to express their indebtedness to Professor C. Spearman and to Professor Rudolf Pintner for many valuable suggestions.

retest was to motivate every pupil to attempt to excel the score of the child who had been his immediate superior in the initial test. This provided an opportunity for the motivation of each pupil, regardless of his ability.

The use of the same form of the test in both initial and final testing periods permits more accurate comparison of the gains than the utilization of different forms. It was felt that the difference in difficulty between the two tests for the groups under comparison introduces unnecessary complications into the analysis of the gains in the final testing. Although this difference in difficulty would probably be negligible in large groups similar to those on which the two forms were standardized, it might nevertheless prove to be considerable in groups that differ from the standard populations.

Procedure

The NIT, Scale B, Form 1 was administered to 42 children. Two equivalent groups were formed, equated for mean and sigma in chronological age and test score. The same form of the NIT was administered again to both groups after an interval of 13 days; in one of the groups, however, the following directions were added:

> "Today you will be given the same test that you had last time in order to let you improve your score. Your standing on the last test was as follows.
>
> (Names were read in rank order)
>
> . first
>
> second, etc.
>
> "A prize will be awarded to each person who gets ahead of the one next above him. For instance Doris B. is just below Harry C. If Doris beats Harry in this test, that is, if she gets a higher score than Harry, she will get a prize, and so will everyone else who beats the one next above him. The pupil who is first now will get a prize if he remains first on this test. Everybody has a chance to win a prize."

Results

There was a reliable gain in score in both the control and experimental groups. The gain in the experimental group (24 children) was slightly greater than the gain in the control group (18 children), but the difference in the gains was only 0.4 times its standard error. The strong incentive of rivalry did not produce a greater gain than the mere repetition of the test under the control condition.

The increase in score is, however, only one of the ways in which the incentive may express itself. The number of items attempted and the number of errors may also be affected. Although there was no difference between the two groups in gain in score, it is possible that differences in gain in the other two measures might have occurred. For this reason the results were analyzed for all three measures—score, number of attempts, and errors.[2]

Both groups tried more items on the retest. The gain in the control group was, however, only two times its sigma, while the gain in the experimental was 7.2 times its sigma. It is apparent that rivalry induced the pupils to exert much more effort than the control motive.

Initially, the control group was somewhat superior to the experimental in the number of items attempted. The difference between the two groups was 4.2 points, which was 1.2 times its sigma. On the retest, however, the experimental group excelled the control by 0.4 points. The net gain was thus in favor of the experimental group, and was equal to 1.8 times its sigma.

Thus, although the experimental group was highly motivated and tried more additional items than the control group on the retest, it failed to excel the control in gain in score. Considering the proportion of correct items to the total number of items attempted, the control group was more accurate than the motivated experimental group.

In the number of errors made, the control and experimental groups show the most marked divergence. The control group decreased its errors on the retest by 2.7 points, equal to two times its sigma, while the experimental group increased its errors by 1.9 points, equal to 1.6 times its sigma.

The control group made a greater number of errors than the experimental in the initial test. The difference was 4.1 points, equal to 1.1 times its sigma. On the retest, however, the experimental group exceeded the number of errors in the control by 0.4 points. The net gain in errors for the experimental group was 4.6 points, which was 2.3 times its sigma.

Comparison of Gains in the Five Subtests. The data for each subtest were analyzed by score, number of attempts, and errors. The results are presented in Tables 1-3. The subtests of the National

---

[2]Ordinarily the difference between the number of items attempted and the score should yield a measure of the number of errors made. Since the various subtests are weighted differently in obtaining the total score, it was considered advisable to compute the number of errors directly, instead of deriving them from the difference between the number of items attempted and the score.

TABLE 1

NUMBER OF ITEMS ATTEMPTED

Test	Control Group Initial M	σ	V	r	Final M	σ	V	Experimental Group Initial M	σ	V	r	Final M	σ	V
1	18.0	2.50	13.9	89	17.9	2.79	15.6	17.2	1.53	8.8	.67	18.9	2.07	11.1
2	36.2	4.29	11.9	41	38.1	2.44	6.3	35.9	3.99	11.1	47	39.0	2.21	5.6
3	38.5	2.63	6.7	52	38.7	2.03	5.1	37.7	2.64	6.8	60	37.9	4.72	12.4
4	29.1	3.19	11.0	41	30.9	1.45	4.8	27.5	4.48	16.1	78	30.3	2.39	8.0
5	32.8	5.09	15.5	.72	33.6	4.51	13.2	32.0	4.88	15.3	82	33.6	5.14	15.0
Total	154.5	9.99	6.5	54	159.2	5.97	3.8	150.3	11.79	7.8	86	159.6	6.30	4.0

TABLE 2

SCORES

Test	Control Group Initial M	σ	V	r	Final M	σ	V	Experimental Group Initial M	σ	V	r	Final M	σ	V
1	29.4	5.69	19.7	92	30.2	4.07	13.7	29.2	4.54	15.5	70	31.2	5.41	17.4
2	25.1	5.40	21.6	97	27.0	5.30	19.6	25.8	5.52	21.2	83	27.4	5.05	18.9
3	23.6	4.37	18.3	53	25.1	4.71	18.8	24.1	5.72	23.8	61	25.3	5.84	23.2
4	20.7	7.19	34.3	87	24.0	5.84	27.6	20.4	6.73	33.5	87	22.5	6.61	28.7
5	29.6	5.34	17.7	57	30.4	4.10	13.7	29.5	4.02	13.3	69	31.6	6.09	19.1
Total	128.7	16.51	12.8	86	136.8	16.69	12.2	129.0	15.88	12.3	87	138.1	20.91	15.1

TABLE 3

NUMBER OF ERRORS

Test	Control Group							Experimental Group						
	Initial				Final			Initial				Final		
	M	σ	V	r	M	σ	V	M	σ	V	r	M	σ	V
1	3.3	1.66	51.5	62	2.7	1.82	66.7	2.6	1.58	61.5	36	3.2	1.69	53.1
2	11.1	4.26	39.1	76	11.1	4.57	41.8	10.0	4.52	45.2	80	11.3	5.48	50.0
3	7.4	2.32	31.1	54	6.8	2.18	32.4	6.8	2.82	41.2	70	6.3	2.97	47.6
4	8.4	6.31	75.0	84	6.9	5.97	87.0	7.0	5.60	80.0	84	7.8	6.10	78.2
5	1.6	1.50	93.8	42	1.6	1.70	106.3	1.3	1.36	104.6	59	1.0	1.17	117.0
Total	31.8	12.22	37.5	85	29.1	13.00	44.8	27.7	11.20	39.3	90	29.5	13.54	46.7

Intelligence Test, Scale B, are as follows 1 Computation Test, 2 Information Test, 3. Vocabulary Test, 4. Analogies Test, 5 Comparison Test

Number of Items Attempted There was a gain in four of the subtests in both the experimental and the control groups. In the remaining subtest, the Arithmetic Test, there was a gain in the experimental group, but a loss in the control group The gain in the experimental group exceeded the gain in the control in each subtest It appears that the incentive caused a gain in the number of items tried on each part of the test.

Number of Errors The difference between the control and experimental group in number of errors was not consistent in the various subtests In the control group there was a decrease of errors in three of the tests, and an increase in the remaining two, while in the experimental group there was an increase of errors in three of the subtests and only a slight decrease in the remaining two, as shown in Table 2.

Score The difference between the control and experimental groups in the gain in score in the various subtests was not as marked as the difference in gain in the number of items attempted. Indeed, the experimental group showed smaller gain than the control in three of the tests This was particularly marked in the Analogies Test The only test in which the gain in the experimental was markedly greater than the gain in the control was the Number Comparison Test, which is definitely a "speed" rather than a "power" test

Effect of Motivation upon Variability. A striking difference was found between the control and experimental groups with regard to variability The standard deviation of the total scores in the control group remained virtually unchanged in the retest, while in the experimental group the standard deviation increased markedly, the gain being equal to 2.7[3] times its standard error The net gain in favor of the experimental group was 1 8 times its sigma With respect to the variability in number of items attempted a similar difference between the experimental and control groups was observed The standard deviation of the number of attempts in the control group decreased in the retest, while the standard deviation of the experi-

[3]The standard error of the difference between the two standard deviations was obtained as follows

$$\sigma_{[\sigma_1 - \sigma_2]} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 - 2r^2_{12}\sigma_1\sigma_2}{2N}}$$

[See Kelley (2, p 178) Formula 121]

TABLE 4

GAINS IN MEAN NUMBER OF ITEMS ATTEMPTED—SCORE AND ERRORS

Test	Group	Number of items attempted Gain	σ_g	G/σ_g	Score Gain	σ_g	G/σ_g	Errors Gain	σ_g	G/σ_g
	Experimental	1.6	0.314	5.42	2.0	0.810	2.47	0.6	0.378	1.70
1	Control	—0.2	0.296	0.06	0.8	0.593	1.34	—0.5	0.361	—1.65
	Difference	1.8	0.4	4.5	1.2	1.0	1.2	1.1	0.5	2.2
	Experimental	3.1	0.723	4.30	1.5	0.646	2.31	1.2	0.670	1.80
2	Control	2.0	0.937	2.04	1.9	0.931	2.04	0.0	0.728	0.00
	Difference	1.1	1.2	0.9	—0.4	1.2	—0.3	1.2	1.0	1.2
	Experimental	0.3	0.788	0.32	1.2	1.050	1.15	—0.5	0.508	—0.09
3	Control	0.2	0.569	0.38	1.6	1.040	1.54	—0.7	0.509	—1.11
	Difference	0.1	1.0	0.1	—0.4	1.1	—0.4	+0.2	0.9	0.2
	Experimental	2.8	0.617	4.73	2.1	0.707	3.00	0.8	0.692	1.15
4	Control	1.8	0.686	2.63	3.3	0.847	3.88	—1.5	3.530	—1.81
	Difference	1.0	0.9	1.1	—1.2	1.1	—1.1	2.3	1.1	2.3
	Experimental	1.6	0.630	2.54	2.2	0.900	2.40	—0.3	0.286	—1.00
5	Control	0.8	0.720	1.10	0.8	1.070	0.77	0.0	0.480	0.00
	Difference	0.8	1.0	0.8	1.4	1.4	1.0	—0.3	0.5	—0.6
	Experimental	9.4	1.287	7.20	9.1	2.040	3.46	1.9	1.240	1.58
Total	Control	4.6	2.300	2.00	8.1	2.000	4.08	—2.7	1.620	—2.00
	Difference	4.8	2.6	1.8	1.0	3.2	0.3	4.6	2.0	2.3

mental group remained practically unchanged. The net gain in favor of the experimental group was 2.2 times its sigma The comparison of coefficients of variation of the scores, also yielded similar results In the control group there was practically no change from the initial to the final test in the coefficient of variation of the scores In the experimental there was an increase in variation. Apparently, the incentive caused an increase in variability in score and prevented a decrease in the variability of the number of attempts [4]

Variation in Subtests. The gain in variability from initial to final testing in the control group is smaller than the corresponding gain in the experimental group in most of the subtests There were, however, marked differences between the individual subtests with regard to the amount of gain In the arithmetic test, the experimental group gained more than the control in the variability of score, but not in the variability of number of attempts The reverse held true in the case of the vocabulary test, the experimental group gained more in the variability of attempts, but not in the variability of scores. A comparison of the coefficients of variation yields similar, though less reliable, results for both of these tests. The experimental group made a greater gain in variability than the control in the Comparison of Numbers Test in both score and number of attempts

An analysis of the coefficients of variation of the individual subtests does not reveal any significant differences for scores and number of attempts In number of errors, however, a striking difference between the two groups is found in Tests 1 and 4 In both of these tests the coefficient of variation of the control group increased in the final test, while that of the experimental group decreased The difference between the gains in the coefficients of variation of the experimental and the control groups is 1 times its sigma in Test 1 and 0.75 times its sigma in Test 2.[5]

[4]In speed tests involving routine work motivation usually decreases the variability of scores. See (3)

[5]The standard error of the difference between two coefficients of variation of two correlated variables is

$$\sigma_{[v_1 - v_2]} = \frac{1}{\sqrt{2N}} \left\{ V_1^2 \left[1 + 2 \left(\frac{V_1}{100} \right)^2 \right] + V_2^2 \left[1 + 2 \left(\frac{V_2}{100} \right)^2 \right] - 2r_{12} V_1 V_2 \left[r_{12} + \frac{2V_1 V_2}{(100)^2} \right] \right\}^{1/2}$$

For explanation of this formula, see (5, p 33)

TABLE 5

GAINS IN STANDARD DEVIATION OF ITEMS ATTEMPTED AND SCORE

Test	Group	Number of items attempted Gain	G/σ_g	Scores Gain	G/σ_g
1	Experimental	0 541	0 03	0 87	1 01
	Control	0 341	0 88	—1 62	—3 13
	Difference	0 20	0 30	2 49	2 8
2	Experimental	—1.781	—3.53	—0 47	—1 15
	Control	—1 851	—2 83	—0 10	—0 11
	Difference	0 07	0 14	—0 37	—0 40
3	Experimental	2 081	3 68	0 12	0 37
	Control	—0 601	—0 15	0 34	0 37
	Difference	2 68	3 9	—0 22	0.20
4	Experimental	—2 091	—4 30	—0 12	—0 17
	Control	—1 741	—3 58	—1 35	—1 70
	Difference	—0 35	—0 5	1 23	1 1
5	Experimental	0 261	0 57	1 95	1 80
	Control	—0 581	—0 93	—2 16	—1 80
	Difference	0 84	1 0	4 11	2 5
Total	Experimental	0 121	0 13	5 60	2 67
	Control	—4 021	—2.60	1 22	0 61
	Difference	3.90	2 2	4 38	1 8

Effect of Motivation upon Correlation between Subtests. The intercorrelations between each of the subtests were computed for the initial and final tests of both the control and experimental groups These are shown in Table 7

There was a gain in the intercorrelations of the subtests between the initial and final testing This held true for both the control and the experimental groups. The gain in the experimental group was 0 82 times its $P.E.$ while the gain in the control was 0 33 times its $P.E.$[6]

It should be noted that generally the intercorrelations between the various subtests are positive, with the exception of four intercorrelations with Test 5.[7] This test is primarily a "speed" test, while the

[6]It should be noted that the standard error utilized in this comparison is probably much larger than the true standard error, since the correlation term between the averages under comparison was omitted

[7]The correlation between the Number Comparison Test and the Information Test was —61 ± 10. This is rather surprising in view of the reputed high intercorrelation among the subtests of the NIT

TABLE 6

GAIN IN COEFFICIENT OF VARIATION IN NUMBER OF ITEMS ATTEMPTED, SCORES, AND ERRORS

Test	Group	Scores Gain	Number of items attempted Gain	Errors Gain
1	Experimental	1.9	2.3	-8.4
	Control	-6.0	1.5	15.2
	Difference	7.9	0.8	-23.6
2	Experimental	-2.3	-5.5	4.8
	Control	-2.0	-5.6	2.7
	Difference	-0.3	0.1	2.1
3	Experimental	-0.6	5.6	6.4
	Control	0.5	-1.6	1.3
	Difference	-1.1	7.2	5.1
4	Experimental	-4.8	-8.1	-1.8
	Control	-6.7	-6.2	12.0
	Difference	1.9	-1.9	-13.8
5	Experimental	5.8	-0.3	12.4
	Control	-4.0	-2.3	12.5
	Difference	9.8	2.0	-0.1
Total	Experimental	2.8	-3.8	7.4
	Control	-0.6	-2.7	7.3
	Difference	3.4	-1.1	0.1

TABLE 7

INTERCORRELATIONS BETWEEN INITIAL TESTS AND BETWEEN FINAL TESTS

Tests	Control Group Initial	Control Group Final	Experimental Group Initial	Experimental Group Final
1-2	2322	2400	.1335	5766
1-3	1476	2644	1596	1382
1-4	5044	5139	1014	5670
1-5	1170	5357	-0930	2880
2-3	2315	2040	5876	3389
2-4	4634	4710	6827	6275
2-5	-6120	-3190	0680	2190
3-4	5694	3764	2992	5123
3-5	0320	1616	0230	1500
4-5	-1010	1555	1450	0780
Averages	1584	2364	2107	3496

others are "power" tests. Excluding Test 5 from the above comparison, the gain in the intercorrelation for the experimental group is much more pronounced. The average intercorrelation of the first four subtests for the initial and final performance for the two groups was as follows·

	Experimental	Control
Initial	3274	3851
Final	4936	3470
D	1662± 1578	0111± 1857
D/PE	1 05	0 05

The difference between the average intercorrelations for the control group is 05 times its $P E.$, while for the experimental group it is 1 05 times its $P E$

The slight increase in the intercorrelations of the motivated tests suggests a stronger bond between two motivated than between two unmotivated tests In order to ascertain whether the motive produced any changes in the factor patterns of the variables, the tetrad difference equations for the initial and for the final testings were obtained in both the control and the experimental group In view of the negative correlations of the "comparison test," it was not included in this comparison

It will be noted that in the control group and in the initial testing of the experimental group the tetrad criterion is satisfied, the difference (largest t over median $P.E$) being virtually zero Upon the introduction of the incentive in the final testing in the experimental group, the tetrad differences become more noticeable This

TABLE 8

TETRAD DIFFERENCES OF INITIAL AND OF FINAL INTERCORRELATIONS

	t_{1234}	t_{1243}	t_{1432}	$P.E$	$\frac{t}{PE}$
Control Group					
Initial	0621	0161	0460	.0851	0 7
Final	—.0310	—.0108	—.0292	0788	—0 4
Experimental Group					
Initial	— 0698	— 0200	— 0498	.0858	—0 8
Final	2076	.1020	1056	0772	2.7

may indicate that incentive conditions introduce a special bond between some of the tests but not between all of them The difference is, however, only 2.7 times its *P E*, and hence not conclusive

Correlation between Initial and Final Tests The correlation between the initial and final scores was approximately the same for the control and experimental groups—about 87 Apparently, the incentive did not disturb the rank order of individuals in the experimental group The correlation between the number of errors in the initial and final testing is also about the same for both groups, being 85 for the control and .90 for the experimental. In the number of items attempted, the correlation is higher for the experimental than for the control group, being 86 for the experimental and only .54 for the control. This difference is more than 3.2 times its *P.E* These correlations are presented in Table 9.

Correlation between Initial Score and Gain In the control group

TABLE 9

CORRELATIONS BETWEEN INITIAL AND FINAL TESTS

Test	Scores Control	Scores Experimental	No of items attempted Control	No of items attempted Experimental	Errors Control	Errors Experimental
1	92	69	89	.67	62	36
2	97	82	41	47	.76	80
3	54	61	52	60	54	70
4	87	87	41	78	84	.84
5	.57	69	72	82	42	59
Total score	87	.86	54	86	85	90
Average (not including total)	77	74	59	67	64	66

TABLE 10

COMPARISON OF GAINS BY INITIAL SCORE

	Control Mean	Control σ	Experimental Mean	Experimental σ	Difference Mean	Difference σ
Above median	6 33	1 76	10 08	1.9	3 75	
Below median	10 44	3 33	7 83	3 6	—2 61	
D	—4 11	3 77	2 25	4 07	6 36	5 51
D/σ_D	—1 09		0 58		1 1	

TABLE 11
COMPARISON OF GAINS BY AGE

	Control M	Control σ_M	Experimental M	Experimental σ_M	D	σ_D
Above median age	7.22	1.41	12.83	1.43	5.61	
Below median age	9.67	1.14	11.58	0.81	1.91	
D	—2.45	1.81	1.25	1.64	3.70	1.62
D/σ_D	—1.39		0.81		2.3	

the correlation is —.29 and in the experimental it is .18 This difference in correlation would indicate that the pupils who made a high score on the initial test gained less than the pupils who made a low score on the initial in the control group, while the reverse holds true for the experimental.

The following table shows the mean gain for those who were above and those who were below the median in the initial test

In the control group the mean gain on the repeated test of the superior children (those who were above the median in the initial test) is larger than the mean gain of the less able children (those that were below the median) and the reverse holds true for the experimental group. In the control group the children who scored high in the initial test seem to have relaxed their efforts in the second application of the test, while in the experimental group the better pupils gained more than the poorer ones.[8]

Number of Items Attempted—Initial and Gain The correlation for the control group is —.57 and for the experimental it is —.35, an increase of .22, equal to 1.5 times its *P.E.* In the control group, the pupils who attempted a large number of items in the initial gained less than those who attempted a smaller number of items in the initial test

Correlation between Age and Gain In the control group the correlation is —.47, while in the experimental group it is .23, a difference which is more than 4 times its *P.E.* It appears that the younger children gained more than the older in the control group and less than the older in the experimental group Table 11 gives the mean gain of the older and younger children in the control and experimental groups respectively.

[8]The factor of age should have been partialled out in this comparison, but on account of the small number of cases, it was considered impracticable

The older children excelled the younger in the experimental group but were inferior to them in the control group There was thus a net gain in favor of the older children of the experimental group equal to 2.3 times its sigma

Summary

1. The repetition of an intelligence test under a very strong incentive of rivalry caused no greater gain in score than repetition under normal conditions

2 The incentive brought about an increase in the number of items attempted There was, however, a corresponding increase in the number of errors, thus resulting in no increase in score.

3. There was a marked difference between the subtests in this respect. On the speed test (Comparison of Numbers) there was an increase in score, while in the power test (Analogies Test) there was a decrease in score under incentive conditions.

4 The intercorrelation between subtests under incentive conditions was slightly higher than under the normal conditions.

5. The correlation between initial and gain under incentive conditions was positive (+ 18), while under the control conditions it was negative (— 29)

6 The correlation between age and gain for the group retested under incentive conditions was positive (+ 23) For the control group it was negative (— 47)

7 The correlation between the number of items attempted on the initial and final tests was higher under incentive conditions than under normal conditions

In conclusion, we may say that the incentive of rivalry failed to bring about a significant increase in score on an intelligence test, although the subjects worked faster when motivated by rivalry The increase in speed resulted only in an increased number of items tried and a corresponding increase in error, leaving the final scores practically unchanged

There was a slightly greater variability of scores under incentive conditions than under normal conditions The various self correlations and intercorrelations were somewhat increased under incentive conditions

The older children increased their score more under the incentive than the younger ones. When the test was repeated without an

added incentive the younger children gained more than the older children.

REFERENCES

1. HURLOCK, E. B. The value of praise and reproof as incentives for children *Arch Psychol*, 1924, **11**, No 71. Pp 78.
2. KELLEY, T L. Statistical method New York Macmillan, 1923 Pp. xi+390
3. MALLER, J B Cooperation and competition *Teach Coll Contrib Educ*, 1929, No 384 Pp 176.
4. THORNDIKE, E L Measurement of intelligence. *Psychol. Rev*, 1924, **31**, 219-252
5. ZUBIN, J Some effects of incentives *Teach Coll Contrib Educ*, 1932, No 532. Pp 60

Teachers College
Columbia University
New York City

L'EFFET DES MOBILES SUR LES RÉSULTATS DES TESTS D'INTELLIGENCE

(Résumé)

On a fait subir le National Intelligence Test, Echelle B, dans les conditions ordinaires standardisées On a formé deux groupes comparables et les a divisés selon l'âge chronologique et l'âge mental. Puis on a répété le test après un intervalle de 13 jours Dans un groupe, l'expérimental, on a annoncé l'ordre d'échelle des sujets sur le premier test et on a promis un prix à chaque sujet qui dépasserait son supérieur immédiat dans le deuxième test Dans le second groupe, celui de contrôle, on a répété le test avec nulle variation de processus.

Le fort stimulant de rivalité n'a pas causé de résultats beaucoup plus élevés dans le National Intelligence Test Le groupe expérimental n'a pas fait de plus grands gains dans les résultats finals que le groupe de contrôle Quand on a considéré le nombre des choses essayées et le nombre des erreurs faites, on a constaté que le groupe expérimental a fait de plus grands gains dans le nombre d'essais. Il y a eu cependant un accroissement correspondant du nombre d'erreurs Comme résultat, les résultats finals sont restés relativement influencés Le stimulant a influencé les tests de "puissance" et ceux de "vitessse" d'une façon différente Dans ceux-ci, les résultats ont montré un vrai accroissement, tandis que dans ceux-là ils ont montré un décroissement

La corrélation moyenne des tests secondaires les uns avec les autres dans les conditions avec stimulant a dépassé celle de ces tests dans des conditions sans stimulant L'analyse des tétrades a montré que les conditions avec stimulant ont résulte dans une différence marquée de tétrade La corrélation entre l'initial et le gain et celle entre l'âge et le gain ont été positive dans le group expérimental, +0,18 et +0,23, respectivement, et négative dans le groupe de contrôle, —0,29 et —0,47, respectivement

La corrélation entre le nombre initial et le nombre final d'essais a été plus élevée dans les conditions avec stimulant que dans celles sans stimulant La variabilité des résultats a été un peu plus élevée dans les conditions avec stimulant que dans celles sans stimulant

MALLER ET ZUBIN

DIE EINWIRKUNG DER MOTIVIERUNG AUF DIE IN INTELLIGENZPRÜFUNGEN ERZIELTEN ZAHLEN

(Referat)

Es wurde die National Intelligence Test, Skala B, unter den gewohnlichen standardizierten Bedingungen gegeben. Es wurden zwei vergleichbare Gruppen gebildet, die in Bezug auf chronologisches und geistiges Alter einander gleichgestellt worden waren Nach einer Zwischenzeit von 13 Tagen wurde die Prufung wiederholt In der einen Gruppe, die die Versuchsgruppe bildete, wurde die an der ersten Prufung erzielte Rangordnung gemeldet, und es wurde jeder Versuchsperson, die in der Wiederprufung ein besseres Resultat erzielen wurde, als die nachst uberlegene Versuchsperson in der ersten Prufung, ein Preis versprochen In der zweiten Gruppe, der Kontrollgruppe, wurde die Prufung ohne Veranderung des Verfahrens wiederholt

Der starke Ansporn des Wetteifers bewirkte in den in dem National Intelligence Test erstelten Zahlen keine bedeutende Erhohung Die Versuchsgruppe erzielte in der endgultigen Zahl keine grosseren Gewinne (gains) als die Kontrollgruppe Wurden die Zahl der versuchten einzelnen Aufgaben (items) und die Zahl der Fehler in Betracht gezogen, so zeigte es sich, dass die Zunahme in der Zahl der versuchten Aufgaben bei der Versuchsgruppe die grossere war, Es bestand aber eine entsprechende Zunahme in der Zahl der Fehler Folglich blieb die endgutige Zahl relativ unverandert Die Prufungen der "Tuchtigkeit" (power) wurden durch die Anspornung anders beeinflusst, als die Prufungen der Schnelligkeit (speed tests). In den ersteren zeigte sich eine bestimmte Verbesserung der Zahl, wahrend in den letzteren eine Verschlimmerung stattfand

Die mittlere Inter-Korrelation der untergeordneten Prufungen (subtests) unter Gegenwart der Anregung war der Inter-Korrelation bei Abwesenheit der Anregung uberlegen Durch vierfaltige (tetrad) Analyse wurde erwiesen, dass die Umstande der Anregung einen ausgepragten vierfaltigen Unterschied (tetrad difference) bewirkten Die Korrelation zwischen der anfanglichen Zahl (initial score) and der Zahlzunahme (gain) und die Korrelation zwischen Alter und Zahlzunahme war in der Versuchsgruppe positiv—bzw + 18 und + 23, und in der Kontrollgruppe negativ—bzw — 29 und — 47

Die Korrelation zwischen der anfanglichen und der endgultigen Zahl der versuchten Aufgaben (attempts) war unter anregenden Umstanden hoher, als unter nicht-anregenden Die Variabilitat der erzielten Zahlen war unter anregenden Umstanden etwas hoher, als unter nicht-anregenden.

MALLER UND ZUBIN

THE DETERMINATION OF A RELIABLE INTELLIGENCE QUOTIENT FOR THE YOUNG CHILD*

From the Iowa Child Welfare Research Station, State University of Iowa

RUTH UPDEGRAFF

This study was undertaken with a view to testing the reliability of an intelligence quotient determined for a child of preschool age just previous to his initial experience in a preschool group Due to the nature of the problem the results should be of significance in indicating the reliability of an initial test given in any new situation by an examiner strange to the child. It was possible, secondarily, to determine the reliability of an initial test given under circumstances to which the child is accustomed and to study the effect of certain variables upon the results

The importance of the early determination of a reliable rating of the intelligence of a school child is widely acknowledged Not only is the rating desirable in dealing with the individual child, but, if determined previous to school entrance, it is of value as a selective aid. On the other hand, an unreliable estimate of a child's level of intelligence is, under the most fortunate circumstances, of no value and, when misleading, it is a decided hindrance. If, therefore, either speed or accuracy need be sacrificed, the latter must remain The problem is that of finding how soon a reliable estimate may be obtained and, if this is a problem with the school child, it is of equal if not greater importance in the case of the child of preschool age for whom the standardization of the test is less perfect and secondary factors more influential

Although the reported findings in studies of the constancy of the intelligence quotient in older children are so mutually consistent that further work would seem unnecessary, the data concerning preschool children are scarce That the IQ's of children of school age are sufficiently constant for the usual predictive purposes has been quite satisfactorily indicated (1, pp 23-53, 3, 4, 5, 6, 7, 9, 10, 12, 14, 15, 16). To assume, however, that the results in these studies would be indicative of the outcome when similar methods

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, December 7, 1931

are applied to the young child would be unjustifiable, the difference in standardization of the test at the preschool and school ages indicates inaccuracy of the testing medium, not to speak of the emotional factors which may have, if not a greater, at least a different influence on the younger child in the test situation. In addition, the present problem involves not only the constancy of the intelligence quotient, regardless of the instability of the testing medium, and of possible real fluctuations in intelligence, but also the possibility of the effect of preschool training upon the intelligence quotient. If preschool training raises the IQ permanently, a significant measure may be unobtainable.

It has been postulated by Johnson (12) that the second intelligence quotients of children tested for the first time at the youngest ages vary from results of the first test more than the quotients of older children, Hildreth (11) indicated that the change is apt to be an increase for children of from three to five and of from six to eight, while for older children the tendency is for the quotient to decline Terman (16, pp. 138-164), citing retests of 99 children of preschool age, felt that age at the time of the first test was not a significant factor in prediction. Baldwin and Stecher (2, p 61), Goodenough (7, 8), and Johnson (12) achieved varying results, claiming prediction to be more or less successful, with anywhere from as many as 17% to as few as 5% of the children changing their IQ by approximately 20 points It is difficult to make comparisons, for not only are the results computed on different bases, but the children vary in age from group to group Different tests were used and intervals between tests differed From the results of testing 43 preschool children twice, at intervals of from 7 to 14 months, Woolley (17) found that only 16% of the children changed 5 or less points in IQ, while 37% changed from 15 points or more These data, so at variance with the earlier data of Terman and Baldwin, she interpreted as indicating that nursery-school training has an effect, though perhaps not a permanent one, upon the intelligence quotient, her conclusions were also based upon a comparison with non-nursery-school children Goodenough (8) failed to find that the intelligence quotients of nursery-school children increased upon retest more than did those of children who had not attended nursery school in the interim Although the average IQ's of both groups increased, that of the nursery-school children more than the non-nursery-school children, nevertheless she did not judge the differ-

ence significant. Although these studies offer only a limited amount of data, and explanation should, therefore, be made with reservations, there are several contributing factors which may have operated to bring about such differing results. Variety of interpretation is also possible In addition to the fact that different revisions of the Binet tests were used, which fact has more weight since the lower age limit of the Minnesota study was 6 months below that of Woolley's, Goodenough's cases were matched as carefully as possible with controls. Each nursery-school child was paired with one like him in sex, age, IQ on first test, interval between test, paternal occupation, education, and nativity of parents It would seem, therefore, that for comparative purposes, the latter study was more rigidly controlled. On the other hand, Goodenough's retests were made after an interval of only 6 weeks, while the Detroit children were examined for the second time at intervals of from 7 to 14 months after the first test. It would be possible to postulate that, although the nursery-school children in Minnesota had perhaps become more familiar with the testing situation in 6 weeks, which could account for any existing superiority, they had not been exposed to the nursery-school situation long enough for the training to have had much effect otherwise. According to Goodenough, this was checked by retesting both groups a second time after about 6 months, there were similar increases for both groups, the control group gaining, although not so much as did the nursery-school group Differences, she asserts, are not real differences in intelligence but are due to imperfect standardization of the tests

In view of the above results, further information is needed It remains to be explained why the tests of children tested for a second time at three years, for example, yield higher intelligence quotients than those of children tested first at three years, for obviously this cannot be laid at the door of imperfect standardization. Is practice the differing factor? Is the child's first IQ simply not a true index because the new situation is inhibitory?

The present study should give further data on these problems By comparing two groups of children receiving second tests in the spring and first tests the preceding fall, one group being tested first before school entrance, the other after the children had become accustomed to the testing situation, it would seem that the emotional factors might at least be delineated. Supposedly, for the second group they would be less than for the first.

For purposes of comparison, 260 children entering the preschool laboratories of the Iowa Child Welfare Research Station for the first time who were, at the time of their first examination, from 19 to 66 months old were grouped as follows:

Group A: Children who received their first examination during the week previous to the opening of the laboratories in late September and who were retested the following spring

Group B· Children who received their first examination after they had attended the laboratories for at least two weeks and before they had attended two months and who were retested the following spring.

Group C Children of Groups A and B who received a third examination during the following year

Group D Children for various reasons not included in Groups A and B but who received a second and third examination within a one-year interval.

The upper and lower limits of the age range of Groups C and D were each raised a year above Groups A and B These children were attending the laboratories at the time of all their examinations with the exception of those in Group A, all of whom started attendance within a week after the examination.

The measures were obtained by the Stanford and Kuhlmann revisions of the Binet-Simon tests Although during the first seven years the laboratories were in session only the former was used, for the last three years the Kuhlmann revision has been given to all children 42 months of age and below Those children, therefore, whose first and second and second and third tests were obtained by different measures, due to an advance from one age group to the next, were not included in this study Not all of the children were given both tests by the same examiner However, since only trained examiners gave the examinations, and since there was a sufficiently large number of them to offset any specific tendency, this fact does not seem to be of importance as a source of error

All tests were given in examining rooms in the laboratory buildings The children in Groups B, C, and D were well acquainted with the examiners and thoroughly accustomed to the testing rooms. During a daily session of the laboratories a child was taken from his group when there seemed to be no abnormal conditions such as ill health or behavior difficulty and when he was not engaged in an all-engrossing activity The children in Group A were, most

of them, coming to the Station for the first time. They came with their parents in order to go through the regular registration routine; they did not know the examiners nor did they know the buildings. In some cases the parent was present at the examination but this was not encouraged and was avoided whenever possible The test was usually begun from 10 to 15 minutes after the child arrived and every attempt was made to make the occasion unhurried and non-exciting

The distribution by age and sex of 186 preschool children examined either before or after entrance and re-examined approximately 6 months later is given below

Age, months	Group A Examined before entrance Boys	Girls	Total	Group B Examined after entrance Boys	Girls	Total
19 to 30	8	9	17	3	4	7
31 to 42	6	8	14	19	22	41
43 to 54	8	9	17	23	17	40
55 to 66	9	6	15	19	16	35
Total	31	32	63	64	59	123

The average intelligence quotients of the four groups are quite comparable:

	Children	Test 1 Mean IQ	S.D.	Test 2 Mean IQ	S D	Test 3 Mean IQ	S D
Group A	63	111.71	14.90	121.28	13.38		
Group B	123	112.17	17.12	121.43	18.00		
Group C	87			121.70	17.00	125.2	16.50
Group D	151			120.80	16.41	124.1	16.38

On the surface, if the increase of a child's second IQ over his first, providing the groups are comparable, were due to greater familiarity with the test situation, one would expect that the first IQ of Group B would average higher than that of Group A, because Group B's familiarity at the time of the first test exceeds that of Group A On the other hand, there may be at least two sets of circumstances operating. One group of influences may affect the reliability of the test by decreasing the variation from the first test to the second, while another may be influencing the intelligence

quotient to increase Familiarity could decrease variability and not increase average IQ

The children in Groups C and D, who received a third test within a year after the second, are distributed by age and sex as follows.

Age, months	Boys	Girls	Total
19 to 30		1*	1
31 to 42	11	17	28
43 to 54	24	38	62
55 to 66	28	12	40
67 to 78	8	12	20
Total	71	80	151

*This child was 2 years and 6 months old

In order to determine the relative degree of correspondence between results of the first and second tests of the same children, the intelligence quotients of Group A, first test, were correlated by the Pearson method with the intelligence quotients of Group A, second test, and compared with Group B, first test, correlated with Group B, second test. Thus the reliability of the intelligence quotient when it is determined before school entrance is compared with the reliability when the quotient is obtained after the child has become accustomed to the school situation This technique for determining reliability is the best available but should not be interpreted too literally In it is the assumption that the IQ is a constant, which may or may not be true, and that rate of growth is regular, there may also be a practice effect which may not operate equally for all children

The correlation of the original and the repeated tests for Group A was $+535 \pm 060$, while the original and repeated tests of Group B have a correlation coefficient of $+837 \pm 017$. This correlation means that the children in Group B maintained their relative positions in both tests to a greater degree than did Group A Since the difference between these correlations is a significant one their meaning is quite apparent, an intelligence quotient first determined before preschool entrance is by no means so reliable as one determined in a period of from two to eight weeks subsequent to preschool entrance

That the intelligence quotient determined by the second test was

TABLE 1

75TH, 50TH, AND 25TH PERCENTILES OF AMOUNT OF DEVIATION OF SECOND IQ FROM FIRST FOR GROUPS A AND B

Age, months	Group	Children	Percentile 75th	50th	25th	Total range	Percentage changing 10 points or more
		First IQ determined before school entrance					
19 to 66	A	63	+18	+7	—1	+33 to —33	42.8
19 to 42	A	31	+21	+7	—7	+33 to —33	54.7
43 to 66	A	32	+19	+9	+3	+33 to — 7	50 0
		First IQ determined after school entrance					
19 to 66	B	123	+16	+8	+1	+38 to —11	56 0
19 to 42	B	48	+19	+9	+1	+38 to —11	50 0
43 to 66	B	75	+15	+8	+2	+36 to — 8	45 3

reliable is indicated by the following coefficients of correlation When the second and third intelligence examinations of Group C (those children of Groups A and B who received a third examination within a year of the second) were correlated, the resulting coefficient was $+ 785 \pm .027$. When Group D was added to Group C, r equals $+.720 \pm .026$.

It will be seen (Table 1) that the deviations in the IQ's of the retests from those of the original tests are larger than those found in previous investigations of older children, are comparable to Woolley's (17) and slightly larger than those found by Goodenough (13).

The fact that approximately one out of every two children either increased or decreased his IQ at least 10 points at his second test indicates the care necessary in the interpretation of a single measure. Evidently (Table 2) the chances of decreasing are less than those of increasing, particularly for children whose initial IQ's are 100 or below. (However, it must be borne in mind that this group of children is a selected group, drawn from a homogeneous population, and that only 14 out of the 36 IQ's below 100 were less than 80) No child whose first IQ was less than 100 tested still lower in his second test From the standpoint of simple increase and decrease, regardless of size, this is the only point of differentiation between the three IQ levels as classified. When the size of the positive and negative changes are compared (Table 3), it seems clear that the increases in IQ are significantly larger than the decreases for both groups.

TABLE 2

PERCENTAGE OF CHILDREN WHOSE IQ'S INCREASED AND DECREASED, BY AGE GROUPS AND BY IQ LEVELS

Age, months	Children	Who increased IQ: Below 100	Who increased IQ: 100 to 119	Who increased IQ: Above 120	Who decreased IQ: Below 100	Who decreased IQ: 100 to 119	Who decreased IQ: Above 120	Who did not change IQ: Below 100	Who did not change IQ: 100 to 119	Who did not change IQ: Above 120
					Tested before entrance					
19 to 42	31	16.1	29.0	19.3	0.0	19.3	16.1	0.0	0.0	0.0
43 to 66	32	31.2	40.6	12.5	0.0	3.1	9.4	0.0	0.0	3.1
					Tested after entrance					
19 to 42	48	16.6	33.3	29.1	0.0	8.3	6.3	0.0	4.2	2.0
43 to 66	75	28.0	22.7	29.3	0.0	9.3	5.3	1.3	2.7	1.3

TABLE 3

MEAN POSITIVE AND NEGATIVE CHANGES IN IQ FROM FIRST TO SECOND TEST

Age, months	Group	Positive change Children	Mean	S.D.	Negative change Children	Mean	S.D.
		Tested before entrance					
19 to 66	A	47	15.3	8.9	15	7.9	7.2
19 to 42	A	20	16.6	9.5	11	10.0	7.9
43 to 66	A	27	14.1	7.8	4	3.5	2.3
		Tested after entrance					
19 to 66	B	98	11.9	8.6	18	4.7	2.9
19 to 42	B	38	14.0	9.3	7	6.1	3.2
43 to 66	B	60	11.5	7.6	11	3.7	2.5

Tables 4 and 5, giving the results of the application of Fisher's formula for computing the significance of differences in cases where the population is small, suggest two factors affecting the size of increase and decrease in IQ. There are between 70 and 80 chances in 100 that the larger changes in IQ, found in the group tested first before preschool entrances, are significantly larger. This holds true for the positive changes of both age groups and for the negative change in the younger children It would seem reasonable, in the light of the other results, to believe that decreases in IQ are smaller for Group B than for Group A, the number of children over

TABLE 4

MEAN POSITIVE AND NEGATIVE CHANGES IN IQ FROM FIRST TO SECOND TEST, GROUPS A AND B COMPARED

Age, months	Group	Children	Mean	S.D. est	P
			Positive changes		
19 to 42	A	20	16.60	2.54	between
19 to 42	B	38	13.97		.3 and .2
43 to 66	A	27	14.07	1.86	between
43 to 66	B	60	11.53		.3 and .2
			Negative changes		
19 to 42	A	11	10.00	3.02	between
19 to 42	B	7	6.14		.3 and .2
43 to 66	A	4	3.75	1.31	not sig-
43 to 66	B	11	3.73		nificant

TABLE 5

MEAN POSITIVE AND NEGATIVE CHANGES IN IQ FROM FIRST TO SECOND TEST, AGES COMPARED FOR GROUPS A AND B

Age, months	Group	Children	Mean	*S D* est	*P*
		Positive changes			
19 to 42	A	20	16 60	2 62	between 4 and 3
43 to 66	A	27	14 07		
19 to 42	B	38	13 97	1 73	between 2 and 1
43 to 66	B	60	11.53		
		Negative changes			
19 to 42	A	11	10 00	3 68	between 2 and 1
43 to 66	A	4	3.75		
19 to 42	B	7	6 14	1 33	between .1 and 05
43 to 66	B	11	3 73		

three and a half in Group A who decreased in IQ was small and its insufficiency is self-evident Until further data are available conclusive proof will have to wait To restate, then, increases in IQ for those children tested after school attendance were not so large as the increase of those first tested before attendance. The first IQ of Group B more closely approximated the second than did that of Group A

That age is a possible factor is indicated in Table 5. The intelligence quotients of the younger children, when they change, increase more than those of the older children and decrease more, as well.

Before more fully interpreting the results, let us summarize them briefly

1 The tests and retests of 63 children between the ages of 19 and 66 months, whose first IQ's were obtained just before school entrance, did not correlate so highly as did tests and retests of 123 children of like age whose first IQ's were determined in from 2 to 8 weeks after school entrance The reliability of the first test of the 63 children was low.

2 The second and third tests of the same and similar children correlated to the same degree as the first and second tests of children first tested after school entrance; they, therefore, correlated more highly than the first and second tests of children first tested before preschool entrance

3. Both groups of children made similar gains in intelligence quotients. The mean IQ's were practically the same, as were the variabilities.

4. The children in both groups varied widely in amount of change of IQ from the first to the second test, there being approximately 15 points IQ between the 25th and 75th percentile of change Fifty per cent of the children changed in IQ 10 points or more.

5 The tendency to increase or decrease did not vary with the level of the original IQ, except for the fact that no child whose initial IQ was less than 100 decreased his intelligence quotient in the second test.

6 Increases in IQ were significantly larger in amount than were decreases

7. There were more increases than decreases

8. The IQ's of children first tested before school entrance less closely approximated the second IQ's than those obtained as the result of tests administered from 2 to 8 weeks after school entrance.

9 Changes in IQ's from the first to the second test were larger for children 19 to 42 months old than for children 43 to 66 months old

That Binet examinations given to children before their entrance to preschool yield questionably reliable IQ's while those resulting from tests given at least two weeks after initial attendance are of higher reliability, is supported by two lines of evidence: (1) There is, under the former conditions, higher correlation between first and second tests, and (2) the size of IQ change from the first to the second test decreased. Not only is this fact of significance in an educational organization but it is important for the clinician, the Binet examination, at best of only indicative value, to be interpreted in light of other information, is for young children a varying measure. Although the conditions of examination for Group B are indicated as more favorable, in general, caution must be used in the application of this fact to the individual case in the assumption that under these conditions a given IQ is reliable. As yet only the group trend is clear and specific application is difficult other than that extreme conservatism must be practiced in the interpretation of a single measure of the IQ

To attempt an explanation of these results may be justifiable but further data must precede final conclusions If children's IQ's are

more reliable after school attendance there are two possible explanations: (1) The conditions of examination are more favorable after a child has been in school than when the whole situation is new to him; therefore, the difference is only apparent and does not indicate a real change in IQ; (2) both measures may be valid and the child's intelligence actually does change after a short attendance.

There are no data to support the second interpretation. The mean IQ and variability of Group B on the first test are practically identical with those of Group A Moreover, two weeks' time in a preschool would not seem that efficacious even to the most ardent enthusiast. It is true that the mean second IQ's are higher than the first; statistically, the differences are not large and the practice effect must be held in mind In the absence of a control group further contention for effect on intelligence is withheld.

On the other hand, it seems reasonable to suppose that certain elements in the preschool environment would contribute to making an IQ obtained there a truer measure of the child's ability than one determined by a stranger in an unfamiliar environment. In the first place, the child is well acquainted with the examiner. Marine (13), in her study of the effect of familiarity with the examiner upon Binet-test performance, came to the conclusion that this factor is not important. It seems possible, however, that the examiner's effect as a part of a new, strange environment is of significance even though under the conditions of her experiment Dr Marine believed a strange examiner not a matter for consideration. The subjects of her study were school children, accustomed not only to the school environment but, in the case of a considerable number of them, to the testing situation. She made no attempt to pair her subjects by any criteria of personality and, although they may have been equated in this respect, it would be more desirable to be reasonably assured of this fact before assuming that conditions were equal for both groups and that familiarity and unfamiliarity were really achieved. In addition, her subjects were older than half of the children in this study and it is pointed out here, as well as elsewhere, that emotional factors are more probably encountered in younger children under these circumstances

In the second place, the child in preschool attendance is familiar with the building and the testing-room He has explored both at his leisure and has developed a feeling of security Quite possibly the fact that his mother is not present or near is contributory, there

is no dependence on her nor sensing of her tenseness in a situation which she may be facing with some apprehension The whole picture of strange examiner, new room and building, a child under emotional strain prepared for he knows not what, is an entirely different one from that of a child going into a familiar room with a friend to play a game Herein the explanation of these results seems to lie

To summarize, it has been found that an intelligence quotient obtained just previous to a young child's first experience in preschool is not reliable This fact is significant as being indicative of the reliability of any initial test given to a young child in a situation strange to him. On the other hand, a test given as soon as two weeks after the beginning of preschool attendance is appreciably more reliable. Many of the children who were tested varied in IQ widely from the first to the second test, 50% of them changing their intelligence quotients by 10 points or more. Increases in IQ were larger than decreases and occurred more often. Until further data are available it seems wise to attribute the difference in reliability to differences in emotional state at the time of the examination It has long been known that the test suffered from imperfect standardization at the lower ages It has been assumed that factors involved in test administration, rather than the test itself, were of extreme importance and their consideration essential. The present study presents evidence on both these points and serves as one more proof for the need of caution in the use and interpretation of the intelligence quotient.

References

1. Baldwin, B. T., & Stecher, L I Mental growth curve of normal and superior children studied by means of consecutive intelligence examinations *Univ Iowa Stud*, *Stud Child Welfare*, 1922, 2, No. 1 Pp 61
2. ——— The psychology of the preschool child. New York Appleton, 1925 Pp 305
3. Broom, E Constancy of the I Q *School & Soc*, 1927, 26, 295-296
4. Cuneo, I., & Terman, L M Stanford-Binet tests of 112 kindergarten children and 77 repeated tests *Ped Sem*, 1918, 25, 414-428.
5. Dickson, V E Mental tests and the classroom teacher Yonkers, N. Y: World Book Co, 1924 Pp xv+231
6. Garrison, S C Fluctuation of intelligence quotient. *School & Soc*, 1921, 13, 647-649
7. Goodenough, F L. The Kuhlmann-Binet tests for children of preschool age. Minneapolis: Univ Minn Press, 1928 Pp 146

8 ——— A preliminary report on the effect of nursery-school training upon the intelligence tests scores of young children *27th Yrbk Nat Soc Stud Educ*, 1928, Pt 1, 361-369

9 GRAY, P L, & MARSDEN, R E. The constancy of the intelligence quotient *Brit J Psychol.* (*Gen Sec*), 1923, **13**, 315-324

10 ——— The constancy of the intelligence quotient—further results *Brit J Psychol* (*Gen Sec.*), 1924, **15**, 169-173

11 HILDRETH, G Stanford-Binet retests of 441 school children. *Ped Sem*, 1926, **33**, 365-386

12 JOHNSON, B J. Mental growth of children New York Dutton, 1925. Pp xix+160

13 MARINE, E L. The effect of familiarity with the examiner upon Stanford-Binet test performance *Teach Coll Contrib Educ*, 1929, No 381 Pp 42.

14. ROOT, W T Two cases showing marked change in I Q *J. Appl Psychol*, 1921, **5**, 156-158

15 RUGG, L S Retests and the constancy of the I Q *J Educ Psychol*, 1925, **16**, 341-343

16 TERMAN, L M The intelligence of school children Boston Houghton Mifflin, 1919 Pp xxvii+317

17 WOOLLEY, H T The validity of standards of mental measurement in young childhood *School & Soc*, 1925, **21**, 476-482

State University of Iowa
Iowa City, Iowa

LA DÉTERMINATION D'UN QUOTIENT D'INTELLIGENCE CONSTANT POUR LE JEUNE ENFANT

(Résumé)

Cette étude présente des données sur la constance relative des premiers tests Binet subis dans des conditions différentes Pour faire des comparaisons, des enfants assistant aux laboratoires prescolaires pour la première fois, âgés à l'époque de leur premier test de dix-neuf à soixante-six mois, ont été groupés ainsi (1) les enfants qui ont subi leurs premiers tests peu de temps avant leur entrée au laboratoire préscolaire (2) les enfants qui ont subi leurs premiers tests après avoir assisté au laboratoire pendant deux à huit semaines. Six mois plus tard on a testé de nouveau les deux groupes d'enfants

Les premiers tests et les deuxièmes des enfants testés pour la première fois avant l'entrée, lesquels ont eté donc dans une situation non familière, ont donne une corrélation de 0,535±0,060, les premiers tests et les deuxièmes des enfants testés après l'entree ont donné une corrélation de 0,837±0,017 La constance des deuxièmes tests a ete indiquée par le fait qu'ils ont donné une corrélation élevée avec des troisièmes tests subis en un an Les deux groupes d'enfants ont gagné également, environ dix points du QI, les variabilités ont eté semblables Cinquante pour cent des enfants ont changé leur QI de dix points ou plus, les changements n'ont pas varié avec le niveau du QI Les accroissements ont été beaucoup plus grands

que les décroissements et ont eu lieu plus souvent Les changements du QI ont eté plus grands pour les enfants âgés de moins de quarante-deux mois que pour les enfants plus âgés Les différences de constance ont eté attribuées aux differences de l'état émotif a l'époque de l'examen, les circonstances de familiarité étant plus favorables

UPDEGRAFF

DIE BESTIMMUNG EINES ZUVERLÄSSIGEN INTELLIGENZQUOTIENTS BEI JUNGEN KINDERN

(Referat)

In dieser Arbeit werden Befunde gegeben uber die relative Zuverlassigkeit von anfanglichen (initial) Binetprufungen die unter verschiedenen Umstanden gegeben wurden Mit dem Zweck der Vergleichung, wurden Kinder, die zum ersten Mal in die Vorschullaboratorien (preschool laboratories) eintraten, und die zur Zeit der ersten Prufung 19 bis 66 Monate alt waren, in folgende zwei Gruppen geteilt (1) Kinder, die ihre ersten Prufungen kurz vor ihrem Eintritt in die Vorschule durchmachten, und (2) Kinder, die ihre ersten Prufungen durchmachten, nachdem sie schon seit 2 bis 8 Wochen die Laboratorien besuchten Beide Kindergruppen wurden 6 Monate spater nochmals gepruft

Die Korrelation zwischen den ersten und zweiten Prufungen an den Kindern, die zum ersten Mal vor dem Vorschuleintritt gepruft wurden, und die sich folglich in einer fremden Umgebung befanden, betrug 535 ± 060, die Korrelation zwischen den ersten und zweiten Prufungen an Kindern, die nach dem Vorschuleintritt gepruft wurden, betrug 837 ± 17 Dass die zweiten Prufungen zuverlassig waren, wurde durch die Tatsache erwiesen, dass sie mit dritten Prufungen, die innerhalb eines Jahres gegeben wurden, eine hohe Korrelation zeigten Die zwei Kindergruppen machten in Bezug auf den Intelligenzquotient ahnliche Gewinne—ungefahr 10 Punkte Auch die Abweichungen waren in den zwei Gruppen ahnlich Funfzig Prozent der Kinder erwiesen eine Anderung von 10 Punkten oder mehr im Intelligenzquotient; die Grosse der Anderungen variirte nicht mit der Hohe der Intelligenzquotiente. Es fanden ofter Zunahmen statt als Abnahmen, und die Zunahmen waren auch bedeutend grosser, als die Abnahmen Anderungen des Intelligenzquotients waren bei Kindern, die junger waren, als 42 Monate, grosser, als bei Kindern die alter waren Unterschiede in Bezug auf Zuverlassigkeit [der Prufungsresultate] werden Unterschieden im Gefuhlszustand (emotional state) zur Zeit der Prufung zugeschrieben, wobei Umstande der Vertrautheit (familiarity) sich als gunstiger erwiesen

UPDEGRAFF

"HABITS" OF THREE SUCCESSIVE NURSERY-SCHOOL GROUPS AND SOME RELATIONS BETWEEN TRAITS*

From the Psychological Laboratory of Purdue University

HARRIET E. O'SHEA

The data of this article[1] have been accumulated from 1927 to 1930 in a nursery school on the Pacific Coast There have been no known influences selecting the applicants to the school differently in any of the three years. A comparison of averages and variabilities of certain traits from year to year and the inter-relationships of traits throughout the three years give tentative answers to certain interesting questions.

THE SCHOOL

The nursery school in which the children were being measured was located at Mills College, California, in its Pre-School Laboratory The college is a privately endowed college for women, enrolling about six hundred students, including graduate students It is located on the edge of an industrial city, Oakland, and is readily accessible from Berkeley, Piedmont, San Francisco, San Leandro, Alameda, and is not far from Palo Alto and other towns on the Peninsula. The Pre-School Laboratory was introduced into Mills College to provide observable material for various college classes and for the student body as a whole.

The school has been limited to an enrollment of twenty children, who may enter between eighteen months and three and one-half years. It is in session in the mornings, and through luncheon. The children are sent back to their homes for afternoon naps. A small tuition, payable by the semester, is charged for each child, and two to four tuition scholarships are maintained each year.

The immediate vicinity of Mills College is not one which would understand or could support a nursery school, most of the neighborhood being filled with workers in nearby automobile factories. A few children only have come from the immediate neighborhood; the

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, August, 17, 1931

[1]The material was originally presented at the meetings of the Pacific Coast Nursery School Association, November 1930

rest have come from Berkeley, Piedmont, downtown Oakland, Alameda, and San Leandro. There has at all times been a waiting list for the school from which selections are made on the basis of several criteria—priority of application, age of the child (younger children being preferred), nearness to the school, and transportation facilities if the family lived at a distance There was little selection with reference to parents' occupation or economic status, since some tuition-free scholarships were always maintained The chief factor of selection, as usual in experimental or progressive schools, seemed to be the intelligence of the parents themselves, those who learned of the nursery school and were interested in it forming a group apparently somewhat above the average of the population

The parents included amongst their number architects, engineers, college professors, musicians, lawyers, politicians, airplane pilots, merchants, clerks, operators of small restaurants, persons in a variety of other occupations, and an occasional unemployed father Some of the mothers were or had been employed at such occupations as nursing, teaching, maid service, operating a switchboard, and so on. A mother or two each year was a widow supporting her children.

The staff of the school consisted of the director of the Pre-School Laboratory, two highly trained supervisors who were the head teachers of the school and who had immediate supervision of student teachers and observers, two trained graduate student assistants, a doctor, a trained nurse, a consulting member of the faculty in orthopedics, a consulting member of the faculty in dietetics, a dietician, a maid, and janitor service

There were constantly in the school practice teachers who had received preliminary training before being allowed to practice, observers from various college classes, visitors from outside of the college, and a few students collecting data for research topics The directions for observers in the nursery school were strictly observed at all times with the result that the large number of visitors to the school had no observable effect upon the children

Habit Inventories

Each fall a group of advanced students, graduate students and occasional senior majors, in connection with a course in the psychology of the preschool child, were permitted to make habit inventories of the children. The Andrus Inventory of Habits From Two to Four (1) was used. The observers had all had a number of courses

in psychology and education before they were allowed to enter this particular course.

Each student inventoried five children and compared and discussed her inventory with that of a co-worker before the final records presented in this study were made. Dr Andrus found that a training of five inventories made an observer relatively reliable and accurate.

The records here presented were accumulated by students who had completed this preliminary training in the inventorying technique. In almost all cases the records used here are, furthermore, an average of the inventories made by two observers during the same days In a few cases where the records are single inventories they are the ninth or the tenth inventories made by the observer in question

The inventorying was done each fall, being completed before Thanksgiving

A scoring formula evolved at Mills College (3), differing from the one utilized by Dr Andrus, was used. This method of scoring has since been adopted by Dr. Andrus in the revised form of the inventory (2).

In place of the maximum scores given by Dr Andrus, the following maximum scores are employed

I	Emotional	224
II	Mental	226
III.	Motor	224
IV	Social-moral	
	Part I	108
	Part II	38

The other material used in calculating the final score is as follows·

T=the total score earned by the child
X=situations listed in the inventory which do not occur during the time the child is observed

The formula for calculating the final score is

$$\text{Final score} = \frac{T}{\text{new maximum} - 2X} \times 100$$

In other respects the inventorying and scoring for each item and for arriving at totals follows Dr Andrus' directions exactly

It will be remembered that the Andrus Inventory provides for a survey of the child's reactions to a rich environment which is familiar

TABLE 1
ANDRUS HABITS FROM TWO TO FOUR (1927-1928)
(Revised scoring)

1	2	3	4	5	6	7
Child	Age (months)	Kuhlman *IQ*	Andrus emotional	Andrus mental	Andrus motor	Andrus social-moral
1	18.5	—	19.5	17.3	15.7	—10.6
2	22.5	—	16.6	14.8	24.6	—19.3
3	26.0	—	22.3	34.3	22.8	+21.5
4	26.0	115	20.1	41.9	42.9	+20.8
5	26.5	133	24.3	51.6	45.6	+24.8
6	29.0	102	13.2	32.2	40.2	+29.5
7	29.5	—	23.0	48.4	59.7	+ 0.7
8	31.0	106	21.2	39.4	41.0	+23.6
9	31.5	105	14.9	56.7	56.6	+46.3
10	34.5	—	21.8	57.8	61.2	+34.5
11	34.5	109	19.3	40.4	49.7	+ 7.0
12	37.5	—	15.3	26.8	31.3	+ 7.0
13	39.5	92	18.0	46.6	35.8	+16.9
14	40.5	111	17.9	38.0	31.2	+23.8
Mean	30.5	109.1	19.1	39.0	39.9	16.2
Standard deviation	6.2	11.1	3.2	12.7	13.5	16.9

TABLE 2
ANDRUS HABITS FROM TWO TO FOUR (1928-1929)
(Revised scoring)

1	2	3	4	5	6	7
Child	Age (months)	Kuhlman *IQ*	Andrus emotional	Andrus mental	Andrus motor	Andrus social-moral
1	23	101	15.7	29.9	47.2	26.2
2	24	—	14.0	23.9	43.6	22.1
3	29	110	21.4	38.0	55.7	37.9
4	31	—	18.3	50.5	61.3	22.2
5	35	—	13.1	50.8	57.7	13.8
6	37	122	24.4	65.6	63.8	51.5
7	38	144	14.7	68.6	84.4	18.3
8	38	109	23.1	55.9	76.0	28.6
9	39	168	20.8	62.4	61.9	31.9
10	39	154	20.2	67.9	78.3	53.6
11	40	109	10.8	48.5	56.5	31.8
12	40	112	21.9	53.3	76.3	41.9
13	41	107	19.0	61.8	72.0	31.4
14	42	124	7.1	42.8	79.4	32.2
Mean	35.4	123.6	17.7	51.4	65.3	31.7
Standard deviation	5.9	20.9	5.1	13.4	12.2	11.1

TABLE 3
ANDRUS HABITS FROM TWO TO FOUR (1929-1930)
(Revised scoring)

1	2	3	4	5	6	7
Child	Age (months)	Kuhlman *IQ*	Andrus emotional	Andrus mental	Andrus motor	Andrus social-moral
1	21	120	22.3	28.8	29.9	8.6
2	22	—	16.4	40.7	52.8	4.9
3	23	104	28.86	31.89	53.06	11.46
4	24.5	100	22.05	27.1	50.5	27.25
5	26.5	106	19.35	37.28	55.66	32.39
6	28.5	113	24.36	40.9	52.21	19.31
7	31	116	20.18	47.67	83.91	32.74
8	32	121	29.41	77.5	73.55	29.7
9	33	145	21.52	54.55	74.53	1.20
10	33	130	20.89	58.7	78.86	31.5
11	35	130	60.0	23.07	58.49	8.03
12	35.5	135	18.07	57.57	69.49	9.17
13	39	—	13.52	58.88	64.16	—10.46
14	43	125	18.35	51.05	75.15	30.65
Mean	31.2	120.4	23.9	46.0	62.3	16.9
Standard deviation	5.2	12.9	10.8	14.8	14.0	13.5

to him—the nursery school. No child's program is changed in any way by the investigation. Furthermore, only those occurrences which the observer actually sees and *records* during the three hours of the inventory can be scored.

SCORES RECORDED

The full enrollment in the school could not, in any of the three years, be inventoried on account of difficulties in the observers' schedules, and on account of absences of children whose inventories were begun but could not be finished It happens that 14 children were inventoried each year, those 14 being a purely random sampling of the school (In the first year of the school the upper age limit was half a year higher and a few inventories were made of older children Those inventories are excluded from the present list.)

Table 1 represents the results of the first year (1927-1928), Table 2, those for the second year (1928-1929), and Table 3, those for the third year (1929-1930) In Table 4 will be found the results from the three years together, arranged according to the chronological age of the children at the time the inventory was made

TABLE 4
ANDRUS HABITS FROM TWO TO FOUR (THREE YEARS COMBINED)
(Revised scoring)

1	2	3	4	5	6	7
Child	Age (months)	Kuhlman *IQ*	Andrus emotional	Andrus mental	Andrus motor	Andrus social-moral
1	18 5	—	19 5	17 3	15 7	—10 6
2	21	120	22 3	28.8	29 9	8 6
3	22	—	16 4	40 7	52 8	4 9
4	22 5	—	16 6	14 8	24 6	—19 3
5	23	101	15 7	29.9	47 2	26 2
6	23	104	28 86	31 89	53 06	11.46
7	24	—	14 0	23 9	43 6	22 1
8	24 5	100	22 05	27.1	50 5	27 25
9	26	—	22.3	34 3	22 8	21 5
10	26	115	20 1	41 9	42 9	20 8
11	26 5	133	24 3	51 6	45.6	24 8
12	26.5	106	19 35	37 28	55 66	32 39
13	28 5	113	24 36	40 9	52 21	19 31
14	29	102	13 2	32 2	40 2	29 5
15	29	110	24 4	38 0	55 7	37 9
16	29 5	—	23 0	48 4	59 7	0 7
17	31 0	106	21 2	39 4	41 0	23 6
18	31	—	18 3	50 5	61 3	22.2
19	31	116	20 18	47.67	83 91	32 74
20	31.5	105	14 9	56 7	56 6	46 3
21	32	121	29 41	77.5	73 55	29 7
22	33	145	21 52	54 55	74 53	1.20
23	33	130	20 89	58 7	78 86	31 5
24	34 5	—	21 8	57 8	61 2	34.5
25	34 5	109	19 3	40 4	49 7	7.0
26	35	—	13.1	50 8	57 7	13 8
27	35	130	60 0	23 07	58 49	8.03
28	35 5	135	18 07	57 57	69 49	9 17
29	37	122	24 4	65 6	63 8	51.5
30	37 5	—	15 3	26.8	31 3	7 0
31	38	144	14 7	68 6	84 4	18 3
32	38	109	23 1	55 9	76 0	28 6
33	39	168	20 8	62.4	61 9	31 9
34	39	154	20 2	67 9	78 3	53 6
35	39	—	13 52	58 88	64 16	—10 46
36	39 5	92	18 0	46 6	35 8	16 9
37	40	109	10 8	48 5	56 5	31 8
38	40	112	21 9	53 3	76 3	41 9
39	40.5	111	17 9	38 0	31 2	23.8
40	41	107	19 0	61 8	72 0	31 4
41	42	124	7 1	42 8	79 4	32 2
42	43	125	18 35	51 05	75 15	30.65
Mean	32 1	118 6	20 2	45 3	55 8	21 6
Standard deviation	6.6	16 9	7 6	14 6	17 4	15 8

There are two children whose records appear twice, child Number 5 in Table 1 is child Number 7 in Table 2, and child Number 6 in Table 1 is child Number 13 in Table 2 It may be noted in passing that the Kuhlman retest rises in both cases for these two children, thereby conforming to the general results obtained by Dr. Goodenough in her retesting of young children

It was, unfortunately, not possible to obtain Kuhlman tests for all of the children inventoried each year, as will be seen in the appropriate tables, but in the year when the fewest Kuhlman tests were given there were 8 out of 14 children so measured In the other years a larger portion was measured These numbers are at least sufficient to give a fair indication of the rate of development of the group of children

Although Dr Andrus does not provide for a total score on the four sections of the inventory, and although it is not clear what such a total score would mean, for the purposes of exploring the results of the inventory each child's total score has been calculated and is presented in Table 5.

The Yearly Composition of the Nursery School

In order to facilitate comparisons, the central tendencies and variabilities of the three groups and of the combined total group are collected in Table 6 The school in the second year had a group of *children shifted somewhat towards its upper chronological limit.* The rate of mental development (Kuhlman IQ) was about the same for the second and third years, both considerably above that of the first year Even the first year, however, was above the average of the population The second year includes a greater variety of rates of development than the other years.

The "emotional" section of the Andrus Inventory does not differentiate between desirable and undersirable emotional traits in the final score The score, therefore, indicates only *excitability* without distinction as to elation and depression, or fear and anger, etc

It will be seen that the three groups are closely alike in excitability, the third group being somewhat more excitable than the other two. The oldest and most rapidly developing group, the second, is the least excitable. The variability of the third group in excitability is much greater than the variability of the other two groups.

In mental scores the three groups follow the order of their Kuhlman mental age—the second group is highest, the third group next,

TABLE 5
TOTAL ANDRUS SCORES (THREE YEARS COMBINED)
(Revised scoring)

Child	Age (months)	Kuhlman MA (months)	Andrus total
1	18.5		41.9
2	21	25.2	89.6
3	22		144.8
4	22.5		37.7
5	23	23.2	119.0
6	23	23.9	125.2
7	24		103.6
8	24.5	24.5	127.0
9	26		100.9
10	26	29.9	125.7
11	26.5	35.3	146.3
12	26.5	28.1	144.8
13	28.5	32.2	136.8
14	29	29.7	115.1
15	29	31.9	156.0
16	29.5		131.8
17	31	32.2	125.2
18	31		152.3
19	31	36.0	184.5
20	31.5	33.1	174.5
21	32	38.7	210.1
22	33	47.9	151.8
23	33	42.9	190.0
24	34.5		175.3
25	34.5	37.6	116.4
26	35		135.4
27	35	45.5	149.6
28	35.5	47.9	154.4
29	37	45.1	205.3
30	37.5		80.4
31	38	54.7	186.0
32	38	41.4	183.6
33	39	65.5	177.0
34	39	60.1	220.0
35	39		126.0
36	39.5	36.3	117.3
37	40	33.6	147.6
38	40	44.8	193.4
39	40.5	45.0	110.9
40	41	43.9	184.2
41	42	52.1	161.5
42	43	53.8	175.4
Mean	32.1	39.42	143.7
Standard deviation	6.6	10.68	40.2

TABLE 6

MEANS AND SIGMAS (ALL CASES, AND BY YEARS)

Year	Age (months)		IQ		Andrus emotional		Andrus mental		Andrus motor		Andrus social-moral		Andrus total	
	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ
All cases	32.1	6.6	118.6	16.9	20.2	7.6	45.3	14.6	55.8	17.4	21.6	15.8	143.7	40.2
1927-28	30.5	6.5	109.1	11.1	19.1	3.2	39.0	12.7	39.9	13.5	16.2	16.9	114.2	38.9
1928-29	35.4	5.9	123.9	20.9	17.7	5.1	51.4	13.4	65.3	12.2	31.7	11.1	151.9	34.0
1929-30	31.2	5.2	120.4	12.9	23.9	10.8	46.0	14.8	62.3	14.0	16.9	13.5	150.7	30.1

and the first group lowest. The spread of scores on the Andrus mental section is very closely alike for the three years Whether the mental scores (39, 46, 51) are to be considered rather similar or quite dissimilar there is hardly enough data to conclude. With the present data the standard deviation of the mean for the first year is 3 4 and of the third year is 4 The difference between the means of the first and third years, 7, is 1 3 times the sigma of the difference, which is 5.2 This, of course, is not a large enough difference to put it outside of the realm of a chance difference The difference in mental score between the first and second years, 12, which is 2.4 times the sigma of the difference ($\sigma_{diff}=4.9$), doubtless represents a real difference between the groups Paired in the other two ways the groups cannot be said to be very different in mental score.

In motor performance the second and third years are very much alike and are considerably higher than the first year. Whether the relatively young first group, which actually associated in the school with children still older than any present in the second and third year, was made lower in its motor score than it otherwise would have been by being occasionally deprived of apparatus is a question. The first and third groups are more alike in chronological age than the second and third groups The third group stands about midway between the other two in mental age.

In social-moral traits the second group stands markedly higher than the first and third, which are identical in score, except that the first group is spread somewhat more than the third

In total score, whatever it may mean, the second and third groups are practically identical, and the first group is much lower

In summary, the second group stands highest in three of the four traits measured by Andrus—and, in standing lowest in the fourth trait, excitability, it may be displaying the best rating of the three groups in that trait.

In emotionality the first and second years are alike, in mental score there is a progression from first to third to second, in motor skill the second and third years are alike, and in social-moral traits the first and third years are alike

In total "habit" score, the two groups which are similar in rate of development, though one is older than the other, and consequently higher in mental age, are nearly identical—the second and third years The first group, which is like the third in chronological age, is considerably lower in total score than the other two groups.

TABLE 7

Year	Age	IQ	Emotional	Mental	Motor	Social-moral	Total
1927-28	3	3	2	3	3	2 5	3
1928-29	1	1	3	1	1	1	1 5
1929-30	2	2	1	2	2	2 5	1 5

TABLE 8

Year	Age	IQ	Emotional	Mental	Motor	Social-moral	Total
1927-28	—	—	—	—	—	—	—
1928-29	+	+	—	+	+	+	+
1929-30	—	+	+	+	+	—	+

In considering the question of whether the nursery school under observation attracted similar groups of children from year to year, the answer in a very general way is that it did if a composite is considered—a composite of age, intelligence, four kinds of "habits," and the combination of those habits.

With respect to separate traits the groups varied from year to year, and, furthermore, two groups alike in one trait were found to be very different in a second trait, and in a second trait were like the group from which they had been very dissimilar in the first trait.

The composition of the three groups may be further examined by putting the parts of Table 6 in terms of the order of magnitude of the central tendencies (Table 7), or in terms of the relation of the central tendencies of each year to that of the three years (Table 8).

Such brief analyses, which tend to obliterate important facts such as the distance of central tendencies from each other, throw into relief differences in general characteristics which may exist between the groups.

Figures 1-5 give the scores on the four sections of the Andrus Inventory and the Andrus totals for the three years As pictured here, the three groups seem to be marked off rather clearly from each other.

In the central tendencies and variabilities given in Table 3, and in the simpler analyses of Table 6, one fact is clear and evident, that in their combinations of amounts of traits no two years were exactly alike This means that a nursery-school teacher enrolling her groups of children each year is encountering not only the expected individual differences among her children in personality profiles, but, evidently, a new total problem from group to group, as well. Could an exact

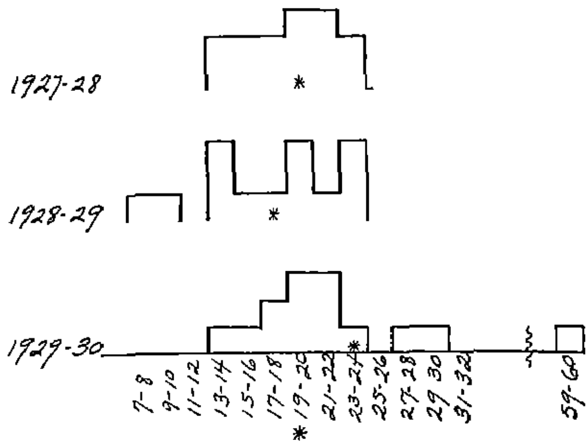


FIGURE 1
EMOTIONAL SCORES
Yearly averages are indicated by stars above the base lines, the total average by a star below the base line

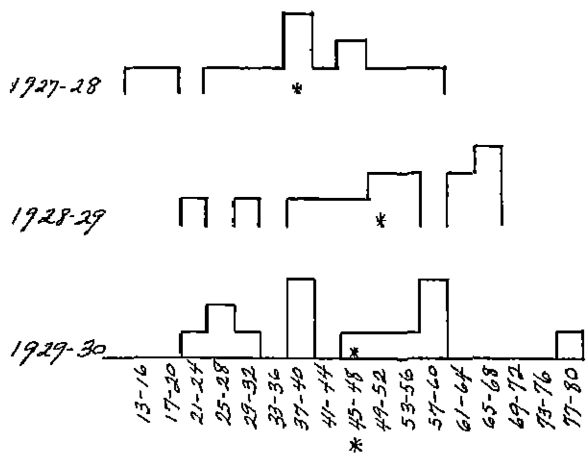


FIGURE 2
MENTAL SCORES
Yearly averages are indicated by stars above the base lines, the total average by a star below the base line

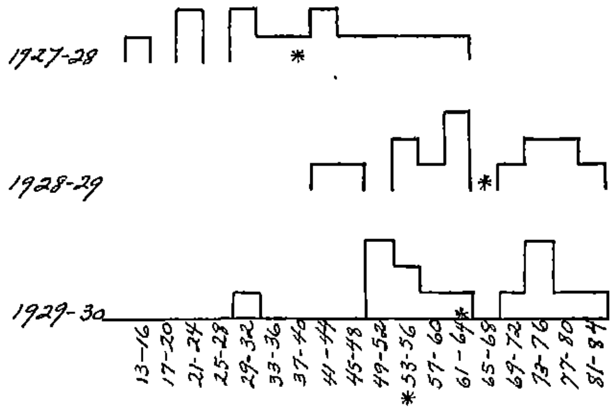


FIGURE 3
MOTOR SCORES

Yearly averages are indicated by stars above the base lines, the total average by a star below the base line.

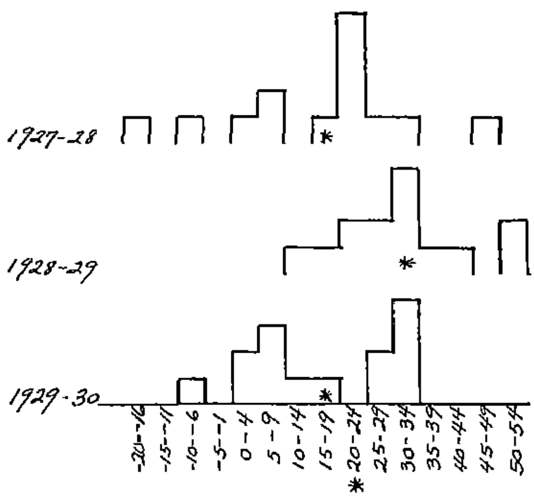


FIGURE 4
SOCIAL-MORAL SCORES

Yearly averages are indicated by stars above the base lines, the total average by a star below the base line

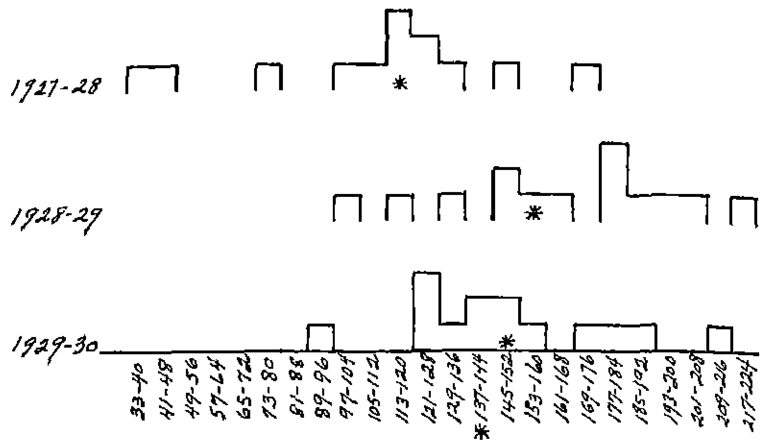


FIGURE 5
TOTAL SCORES
Yearly averages are indicated by stars above the base lines, the total average by a star below the base line

and complete record of the educational situation and the teacher's activities in a school be kept from year to year, apparently it would show, within the similarities established by sound procedure, unique distributions of concentrations of effort and emphasis each year in so far as the teacher is responding expertly to the needs of her group.

AGE GROUPS

If such an instrument as the Andrus Inventory is to be of maximum value, either in diagnosing children or in estimating changes during the year in order to evaluate procedures (assuming for the moment that it has been revised to the point of a sound reliability), it is necessary to establish norms by age, or by age and economic status, or by any other grouping that seems useful

The number of cases in the present study is too small to be of much use in such an undertaking, however, the children have been grouped by four-month intervals, and the central tendencies of the scores are presented in Table 9 for whatever suggestive value they may have.

The emotional scores ("excitability") seem to have no progression of any sort with age They begin at 18 in the youngest group and end at 16 in the oldest group, with 20's in between

TABLE 9
ANDRUS AVERAGE BY FOUR-MONTH PERIODS

Age in months	*N*	Kuhlman IQ (*N*)	Andrus emotional	Andrus mental	Andrus motor	Andrus social-moral	Andrus total
19-22 (18 5)	4		18 7	25 4	30 7	4 1	78 5
23-26	8	109.8 (6)	20.7	34 7	45 1	23 3	124 0
27-30	4	108 3 (3)	21 2	34 8	51.9	21 8	134 9
31-34	9	118 8 (7)	20 8	53 7	64 5	32.0	164 4
35-38	7	128 0 (5)	24 1	49 7	63 0	31 2	156 3
39-42	9	122 1 (8)	16 5	53 3	61 7	28 1	159 7

The mental score rises with jerks It seems to have three levels—at 19-22 months, at 23-30 months, and at 31-42 months.

The motor score rises from 19-22 months to 23-26 months, to 27-30 months, and to 31-34 months, after which it remains constant.

The social-moral score seems also to have three levels. There is a big jump from 19-22 months to 23-30 months, and then a smaller rise to 31-42 months.

The total score rises regularly for the first four age groups, and then seems to remain constant

RELATIONS BETWEEN TRAITS

The number of children in the total group is too small to give reliable correlations, and yet interesting suggestions of relationships between traits can be derived from the present data

Zero-order correlations and various partial correlations may be found in Tables 10A, 10B, and 11 The group is necessarily limited to 31 cases whenever mental age is included in the calculations.

The largest zero-order correlation is that between CA and MA (.83) The correlations of motor score with MA (60) and with mental score (64) are large The correlation of total score with CA (.56) and with MA (62) are also relatively large The largest

TABLE 10A
INTERCORRELATIONS (42 Cases)

	MA	Emotional	Mental	Motor	Social-moral	Total
CA	—	—10	61	56	36	56
MA		—	—	—	—	—
Emotional			—.10	03	—04	—
Mental				74	.33	—
Motor					31	—

TABLE 10B
INTERCORRELATIONS (31 Cases)

	MA	Emotional	Mental	Motor	Social-moral	Total
CA	.83	—17	.53	50	.27	56
MA		—.05	42	.60	21	62
Emotional			—22	—03	—28	—
Mental				.64	41	—
Motor					30	—

negative correlations are those of emotional score with mental score (—.22) and with social-moral score (—.28). The correlation between CA and emotional score shows a distinct negative trend (—.17) The correlations closest to zero are those between MA and emotional score (—.05) and between emotional score and motor score (—03)

TABLE 11
PARTIAL CORRELATIONS (31 Cases)

I *CA constant*	
MA with emotional	16
" " mental	—05
" " motor	.39
" " social-moral	—04
" " total	.33
Emotional with mental	—15
" " motor	07
" " social-moral	—.24
Mental with motor	70
" " social-moral	32
Motor with social-moral	20

TABLE 11 (*continued*)

II. *MA constant*	
CA with emotional	—23
" " mental	36
" " motor	—0002
" " social-moral	.18
" " total	.11
Emotional with mental	22
" " motor	—003
" " social-moral	—.27
Mental with motor	53
" " social-moral	36
Motor with social-moral	.23
III *Emotional score constant*	
CA with social-moral	235
MA with motor	60
" " social-moral	.20
Mental with motor	65
" " social-moral	37
Motor with social-moral	30
IV *Mental score constant*	
CA with emotional	—06
" " motor	25
" " social-moral	068
MA with emotional	048
" " motor	.48
" " social-moral	082
Emotional with motor	14
" " social-moral	—21
Motor with social-moral	06
V *Motor score constant*	
CA with emotional	—18
" " social-moral	145
MA with emotional	—04
" " social-moral	079
Emotional with mental	—26
" " social-moral	—28
Mental with social-moral	297
VI *Social-moral score constant*	
CA with emotional	—10
MA with emotional	—12
" " motor	58
Emotional with mental	0009
" " motor	059
Mental with motor	64
VII. *Total score constant*	
CA with MA	75

TABLE 11 (*continued*)

VIII.	*CA and MA constant*	
	Emotional with social-moral	—23
IX.	*CA and emotional score constant*	
	MA with social-moral	003
X	*CA and social-moral score constant*	
	MA with emotional	09
XI	*MA and emotional score constant*	
	CA with social-moral	.13
XII	*MA and social-moral score constant*	
	CA with emotional	—19
XIII.	*Emotional score and social-moral score constant*	
	MA with motor	58
	Mental " "	.61
XIV.	*CA, emotional score, motor score constant*	
	MA with social-moral	—09
XV.	*CA, motor score, social-moral score constant*	
	MA with emotional	.15
XVI.	*MA, emotional score, motor score constant*	
	CA with social-moral	14
XVII	*MA, motor score, social-moral score constant*	
	CA with emotional	15

It is interesting to see what some of these relationships become as they are analyzed further. Removing the influence of MA, the correlation of total score with CA drops to .11, and, with the influence of CA removed, the correlation between MA and total score drops to .33 The total score, in other words, seems to increase with a combination of increasing MA and length of life, but follows more closely the increase in mental age.

The high correlation between motor score and both MA and mental score is noticeable. It correlates fairly highly also with CA (.50) When MA is made constant, the correlation between CA and motor score drops to zero (— 0002) As was pointed out above, emotional score and motor score correlate zero. Motor score and social-moral score correlate 30.

TABLE 12

PARTIAL CORRELATIONS GROUPED FOR MOTOR SCORES

CA constant	
Motor with MA	39
" " mental	70
MA constant	
Motor with CA	—.0002
" " emotional	—003
" " mental	53
" " social-moral	23
Emotional constant	
Motor with MA	60
" " mental	.65
Mental constant	
Motor with CA	25
" " MA	48
" " *emotional*	14
" " social-moral	06
Social-moral constant	
Motor with MA	58
" " mental	64
Emotional and social-moral constant	
Motor with MA	.58
" " mental	61

Further relations are shown in Table 12 The evidence seems to show clearly that motor score increases as both MA and mental score increase (MA and mental score correlate .42 with each other, but with CA constant, only —.05). Furthermore, excitability and motor skill do not seem related (except that with mental score constant they correlate slightly, 14). Motor score and social adjustments maintain some relationship even with MA constant (.23) or with CA constant (.20), but when mental score is made constant they have a zero correlation (06).

The raw correlation between mental score and social moral score is .41; with CA constant it becomes .32, with MA constant, .36.

The correlation, then, between social adjustments and motor score seems to depend upon the third factor of mental score, which maintains throughout a real correlation with both social adjustment and motor score. The "mental" section in the Andrus Inventory contains a large measure of language control Perhaps language helps

noticeably in social adjustments; and perhaps motor performance and language play back and forth on each other, or perhaps both are subservient to some fundamental maturation phenomenon.

It is to be remembered that the zero-order correlation between MA and mental score (.42) drops to zero with CA constant. With MA constant, CA and mental score correlate .36. The "mental" section of the inventory is, then, measuring something which increases as one lives longer, regardless of one's MA level.

Motor score and mental score do not behave alike with reference to length of life. Although the zero-order correlations with CA are almost identical (.50 and .53), with MA constant, as has been said, the CA–mental score correlation maintains itself fairly well (.36), but the CA–motor score correlation drops to zero (—.0002).

There seems here to be evidence that motor development depends upon a growth factor (correlation with MA rather than length of life) and upon an experience factor (correlation with mental score and social-moral score which seem to have no connection with maturation as evidenced by MA and considerable connection with length of life).

(The zero-order correlations between social-moral score and CA and MA are .27 and .21. With MA constant, the social-moral–CA correlation is .18, and with CA constant, the MA–social-moral correlation is zero, —.04.)

The emotional scores (emotionality or excitability) are associated with the other scores as shown in Table 13. The zero-order correlations of emotionality with the other traits are all negative, all small, and two are practically zero—MA and motor.

The correlation with social-moral score remains negative and in the 20's no matter what factors are held constant. The child who is excitable and the child who is well adjusted socially tend not to be the same child.

(It should be remembered here that the inventory subtracts undesirable from desirable social traits, whereas in the emotional score it lumps all excitability together without attempting to distinguish desirable from undesirable behavior. The material of the two sections is similar in some limited respects. How these circumstances may affect the correlation it is hard to tell.)

As has been said, the correlatoin of emotionality with MA is practically zero. However, when CA is kept constant a positive correlation appears, although it remains small (.16). It would seem then

TABLE 13
CORRELATIONS OF EMOTIONAL SCORES WITH OTHER SCORES

Emotional score with:	Zero order	CA constant	MA constant	Mental constant	Motor constant	Social-moral constant
CA	—17		—.23	—06	—.18	—10
MA	—.05	16		048	—04	—12
Mental	—22	—15	—22		—.26	.001
Motor	—03	07	—003	14		06
Social-moral	—28	—24	—27	—21	—28	

CA and MA constant	
Emotional with social-moral	—.23
MA and social-moral constant	
Emotional with CA	—19
MA, motor, and social-moral constant	
Emotional with CA	—15
CA and social-moral constant	
Emotional with MA	09
CA, motor, and social-moral constant	
Emotional with MA	.15

that there is a little tendency when children are the same age for those who are at a higher mental level to be somewhat more excitable

When the children are all at the same mental level the correlation between emotionality and CA becomes a larger negative quantity (—.23), in other words, the older children are then seen to be the steadier children

The mental score maintains an appreciable low negative correlation with emotionality except that it becomes zero when the children are all equally well adjusted socially. In other words, what the child knows and has gained in mental skills (as tested by the mental section) has no connection with emotionality when the group are all equally well adjusted socially, but when they are alike in other traits taken one at a time the child who knows more is steadier Especially when the children are alike in motor skills, then the child who is better equipped mentally is less excitable.

Motor skill or its opposite seems to have no connection with excitability (all partials zero) except that when the children all know the same amount (mental score) there is some tendency for the more skillful children to be more excitable (r= 14). Otherwise there

does not seem to be any tendency for the clumsier child to be excitable, nor for the excitable child to be skillful

When the children are equally well adjusted socially neither CA nor MA has much connection with emotionality but there does remain a low negative correlation with both factors (— 10 and — 12) which suggests that the older and brighter children may be steadier. However, since with both CA and social-moral constant the correlation with MA is .09, it may be more advisable to remember that within a given age group (exactly the same age) the brighter children tend to be more excitable.

The relationship of social-moral score with other traits is shown in Table 14. Social adjustment has a low positive correlation with all the other traits tested except emotionality. As has been noted the negative correlation with emotionality maintains itself throughout all first-order partials.

When CA is kept constant the correlation between MA and social adjustment is in effect zero (— 04). When MA is kept constant the correlation between CA and social adjustment remains positive

TABLE 14
CORRELATION OF SOCIAL-MORAL SCORES WITH OTHER SCORES

Social-moral score with	Zero order	CA constant	MA constant	Emotional constant	Mental constant	Motor constant
CA	27		18	235	068	145
MA	21	— 04		.20	08	.08
Emotional	— 28	—.24	— 27		—.21	—.28
Mental	41	.32	36	.37		.297
Motor	.30	20	23	30	.06	

CA and MA constant	
Social-moral with emotional	—.23
MA and emotional constant	
Social-moral with CA	.13
MA, emotional, and motor constant	
Social-moral with CA	.14
CA and emotional constant	
Social-moral with MA	.003
CA, emotional, and motor constant	
Social-moral with MA	— 09

and appreciable, though low (.18). Within the age limits measured, and under nursery-school conditions, being socially well adjusted seems to be connected with having lived longer and seems not to be connected with mental level.

The highest group of correlations with social adjustment are those with mental score. In the zero-order correlation and throughout all of the first-order partial correlations the figures are maintained between .30 and .40. Both social adjustment and one's mental skills seem more closely connected with length of life than with MA level, as though they were both learned traits (subject, of course, to the learning capacity of the given mental age).

Whether either facilitates the other, as, for instance, whether knowing more language helps one's adjustments, or possibly becoming better adjusted increases one's perceptual patterns and thus one's language, or whether both are simply independently results of longer life, the present data cannot determine.

When mental score is kept constant the correlations between social adjustment and CA, MA, and motor score all become practically zero (.07, .08, .06). In all the other first-order partials the correlations between social adjustment and motor score have maintained a low positive value (from .20 to .30). Apparently the association of social adjustment with motor score is by way of amount known (mental score) as was suggested above or possibly all three are independently associated with length of life.

When all of the children are alike in motor skills, the correlation of social-moral score with MA level approaches zero (.08), but there is a slight tendency for the older children to be better adjusted (.145) and a greater tendency for those who have more mental skills to be better adjusted (.297).

Comments

Since the number of cases used in the experiment is small, the results can be only suggestive. However, they may serve to raise interesting questions of relationships between traits for further investigation.

Whatever relationships may be found in this and other experiments bring in their train further kinds of questions as to whether either emotionality or social adjustment can be influenced through modifying the other trait, whether devising environments and procedures to increase motor skills will accelerate gain in other desirable traits, and so on.

Possibly experiments on procedure (in its largest sense) and its results in the development of children are among the most interesting and the most significant, even though among the most difficult, psychological experiments waiting to be done.

References

1. ANDRUS, R A tentative inventory of the habits of children from two to four years of age *Teach Coll Contrib Educ*, 1924, No 160 Pp 50
2. ——— An inventory of the habits of children from two to five years of age. *Teach Coll Contrib. Educ.*, 1928. Pp 51
3. O'SHEA, H E Combinations of traits in young children In *The pre-school child: his development and his guidance* *Mills Coll Bull.*, 1928, 29

Purdue University
Lafayette, Indiana

LES "HABITUDES" DE TROIS GROUPES SUCCESSIFS D'UNE ÉCOLE MATERNELLE

(Résumé)

On a fait une liste des "habitudes" (Andrus) de 14 enfants dans chacun de trois groupes successifs d'une école maternelle (Q I. moyens Kuhlman de 109, de 123 et de 129) Les trois groupes ont différé en chacun des quatre traits (émotif, mental, moteur et social-moral) et en leur combinaison de traits, ce qui prouve qu'une maîtresse de la même école maternelle n'a pas la même tâche totale d'enseignement d'année en année

Le moyenne fait à intervalles de quatre mois ne montre aucun changement du résultat émotif Les autres traits et le résultat total s'accroissent jusqu'à 31-34 mois. Le resultat total s'accroît avec l'âge mental et la longueur de la vie mais est plus associé à l'âge mental

Les corrélations entre le résultat moteur et A M (0,60) et entre le résultat moteur et le résultat mental (0,64) sont les plus élevées Quand l'âge mental est constant, l'âge chonologique et le résultat moteur donnent une corrélation de nulle. La corrélation entre le résultat émotif et le résultat moteur est nulle Une corrélation entre l'adaptation sociale et le résultat moteur dépend, paraît-il, d'un troisième facteur, le résultat mental, (que mesure une combinaison de l'âge mental et de l'apprentissage). Le développement moteur dépend et d'un facteur de croissance et un facteur d'expérience Quand l'A C est constant, les enfants au niveau mental le plus élevé tendent à être un peu plus enclins à s'émouvoir Quand l'A.M est constant, il paraît que les enfants les plus âgés sont les moins enclins à s'émouvoir. Le résultat mental a une corrélation négative peu élevée avec l'émotionalité, sauf que la corrélation est nulle quand le résultat social-moral est constant Quand les autres traits, considérés séparément, sont constants, l'enfant qui sait le plus est moins enclin à s'émouvoir.

O'SHEA

GEWOHNHEITEN VON DREI AUFEINANDERFOLGENDEN KINDERGRUPPEN FÜR VORSCHULPFLICHTIGE KINDER

(Referat)

Die Verfasserin notierte die "Gewohnheiten" ("Habits" nach Andrus) *von 14 Kindern aus drei aufeinanderfolgenden Kindergruppen fur vorschulpflichtige Kinder* In den drei Gruppen war der durchschnittliche *Intelligenzquotient nach Kuhlman bezwecklich 109, 123, und 129* Die drei Gruppen waren in Bezug auf alle vier Eigenschaftsgruppen, d.h., affektiv, geistig, motorisch, und sozial-moralisch, wie auch in den Verbindungen dieser Eigenschaften, verschieden Hierdurch wird erwiesen, das der Lehrerin einer Klasse fur vorschulpflichtige Kinder in der selben Schule jedes Jahr eine andere gesammte Lehraufgabe bevorsteht

Nimmt man fur je vier Monate zusammen die Durchschnittszahl, so zeigt sich kein Unterschied in den an den affektiven Prufungen erhaltenen Zahlen (*emotional* scores) Die fur die anderen Eigenschaften erhaltenen Zahlen und die Gesammtzahlen nehmen bis zu 31 bis 34 Monaten zu Die Gesammtzahl nimmt mit Zunahme des geistigen Alters (mental age) und des Lebensalters zu, steht aber zu dem geistigen Alter in einem naheren Verhaltniss.

Die hochste Korrelationen sind die zwischen der an motorischen Prufungen erhaltenen Zahl und dem geistigen Alter (.60) und die zwischen dieser motorischen Zahl (motor score) und der geistigen Zahl (mental score) (.64) Wird das geistige Alter konstant gehalten, so ist die Korrelation zwischen dem chronologischen Alter und der motorischen Zahl Nul Die Korrelation zwischen der affektive Zahl (emotional score) und der motorischen Zahl ist Nul Eine Korrelation zwischen der sozialen Anpassungsfahigkeit (social adjustment) und der motorischen Zahl scheint von einem dritten Bestandteil, der geistigen Zahl abhangig zu sein, die eine Kombination aus geistigem Alter und Lernfahigkeit misst Die motorische Entwicklung ist sowohl von einem Erfahrungsfaktor abhangig. Wird das chronologische Alter konstand gehalten, so zeigen sich die alteren Kinder als die standhafteren (steadier) Die geistige Zahl erweist eine niedrige negative Korrelation mit der Affektivitat (emotionality) ausserhalb der Tatsache, dass wenn die sozial-moralische Zahl konstant gehalten wird, die Korrelation Nul ist Wenn die ubrigen Eigenschaften der Reihe nach konstant gehalten werden, zeigt es sich, dass das Kind das mehr weiss weniger erregbar ist.

O'SHEA

GESTALT PRINCIPLES IN THE SIDEWALK DRAWINGS AND GAMES OF CHILDREN*

From the Psychopathic Department of the Bellevue Hospital

LAURETTA BENDER

Children's drawings have been studied extensively from many viewpoints, and these studies have thrown much light on the conceptional and intellectual development of the child Goodenough (8) has been able to standardize children's drawings of men and thereby offers a simple and apparently satisfactory test for the child's intellectual level of development. Before her, however, Sully (24) in 1897 had made a very valuable study of children's drawings at different developmental stages and arrived at conclusions very similar to those established by Goodenough's more systematized study He outlined three stages: (1) There is an early stage of vague, formless scribblings resulting from an aimless swaying of the pencil to and fro and producing a chaos of slightly curved lines These scribblings are purely spontaneous and have no resemblance to any model presented. They may accidently resemble some form and be given a name, or the child may make believe that they have some meaning which is arbitrarily given, as when Preyer's (18) child in his second year would say that he was "writing houses" while he was scribbling in this fashion. (2) There is a second stage of primitive circular design best characterized by lunar schemes of the human face. This is gradually evolved from the first stage as the curved lines become loops, and crudely placed dots indicate that the relationship of feature is more important than their details. Trunk and limbs are entirely unimportant (Plate 6, Figure 1) (3) In the later stage there is a more sophisticated treatment of the human and animal forms. Sully makes the interesting observation that the early pictorial forms of human beings and animals have an embryonic configuration. He concludes that the child as an artist is more of a symbolist than a naturalist

*Recommended by Paul Schilder, accepted for publication by Carl Murchison of the Editorial Board, and received in the Editorial Office, January 28, 1932

The report of investigations the world over shows a remarkable constancy in the order of development both in regard to the method of indicating separate items and their relationship to each other and the order in which these items and relationships tend to appear Rouma (19) who had made one of the most extensive studies, including the productions of subnormal children, found that the drawings of these children were similar to those of younger children and were characterized by frequent regressions to the inferior stages with special emphasis on automatisms, flight of ideas, fragmentation, meticulousness, and frequent recurrences of the same form

The careful studies of individual children probably give us as much information as any other approach to the subject. Baldwin (2) followed the course of his child in her ability to draw from the 19th to the 27th month At the earliest age there was "the simplest and vaguest and most general imitation of the teacher's movements, not the tracing of the mental picture There was no semblance of conformity between the child's drawing and the copy She could not identify it herself." He found that at the 19th month there were only sweeping, whole-arm movements from the shoulder, a few months later she began to flex the elbow and wrist, and at 27 months she was able to manipulate her fingers. At that time her productions were made of loops in the clockwise direction with an emphasis on the horizontal plane. In each new drawing there was a tendency to carry over a whole or a part of the previous drawing Thus it would appear that the motor functions develop in part separate from the perceptive functions Similar observations were made by Preyer (18) in the development of his son. Ament (1) found that at three years of age there is no resemblance between the child's drawing and the model Later, the first recognizable drawings were of man made up of loop units and then of animals made in the same way. Some of Ament's illustrations of children's drawings of animals are of particular interest. They were made in such a way that there was a long horizontal loop for the body and a small round loop at the left side for the head, but the body loop extending to the right was bent around a little and was terminated by a curling tail that bent with the whole figure in a clockwise direction. Near the head was the first small loop for a leg and this was perseverated five or six times without regard for the natural number of legs, and all were bent with the figure in a clockwise direction so that the whole gave the marked impression of a whirl or vortical design In the same way the child

made a stork, cow, pig, dog, mouse, and cat Shinn (22), in following the development of her niece, found that the child could recognize *O* in her 12th and 13th month, although she then tended to confuse it with *G* and *Q*, which, however, she learned to distinguish in the next few weeks. In her 109th week she drew circles in imitation of her aunt's arm movements and made many spontaneous scribblings back and forth over each other.

Goodenough, from a careful analysis of published studies and her own material, made something of the following summary In young children there is a relationship between concept development as shown in drawings and in general intelligence. Drawing is primarily a language or form of expression to the child. In the beginning he draws what he knows This is Verworn's (25) ideoplastic stage. Later he attempts to draw what he sees This is the physioplastic stage The features of the ideoplastic stage are seen in the child's tendency to exaggerate size of items that seem important or interesting, and to minimize or omit the other parts The child pays little attention to the details of the object before him. Goodenough's psychological interpretation of children's drawings is summarized as follows. Association by similarity; analysis and evaluation of component parts and their spatial relationships and proportions; judgment of the same; abstraction, reduction, and simplification of parts, coordination of eye and hand movements, adaptability These conclusions were influenced in part by those of Paulsson (17), where he concludes that "meaning" is the guide to graphic structure in their development from primitive schematic stages to the highest manifestations of the artistic mind Such conclusions were based on his studies of adults' efforts to interpret meaningless blots into beautiful drawings according to the interests of the subjects From this material he concluded that association by similarity is an important part in graphic art and that the first impulse towards graphic expression has its origin in the desire for emotional outlet and pleasure derived from objectification of emotion A much earlier analysis by Cushman (7) seems to give us more insight into the subject; she says, "All the fine arts have this in common—they interpret the human mind. This mind is so constituted that it seeks organization of the material with which it deals. Therefore logical unity or arrangement is the basic principle of all art In those arts which appeal through the eye the unity must be spatial This unity is more important than detail." Coming in 1908, this is an interesting anticipation of the

Gestalt theory of sensory organizaton, although, of course, Kohler (13) goes much further, stating that "order and distribution in the sensory field is in each case the result of sensory dynamical interactions," and "that all experienced order in space is a true representation of corresponding order in the underlying dynamical context of the physiological process"

The classical statement that the "child draws what he knows, not what he sees" seems to have had undue influence on the psychological interpretation of children's drawings It is obviously unfair to the child He knows a great deal more than he draws. Thus, Goodenough herself points out the fact that the small child can recognize his mother or even a small photograph of his mother, although neither he, nor, indeed, most adults, could tell by what detail or combinations of details this is possible, nor could he or most adults draw a picture of the mother, either from life or by copying a photograph, by which he or anyone else could later recognize the mother. Or, again, as Goodenough points out, a child of three can point to his own hair or that in a good picture, but he almost invariably omits the hair when he draws a picture of a man.

It thus becomes apparent that what a child knows is based on perceptional and therefore conceptional recognition of certain constellations or *Gestalten* that have significance for him, and that this capacity is not based upon his recognition or analysis or evaluation of details as such, but of their relationship to each other and to his whole world On the other hand, his graphic representation of his world is based on his ability to make symbols that represent these relationships of parts and their total meaning to him This ability is one that, even in the average adult, falls far below what he knows about his world Our problem, therefore, is to discover the mechanism for the origins of graphic symbols.

Of course, in such a problem there are both motor and sensory elements which are so interwoven that they cannot be separated and which form an organized perceptional-motor pattern. Koffka (14) has pointed out that even during development all motor acquisitions have a sensory component, and it is also held that movement is a necessary condition for perception, at least in the primitive stage of development (Katz, 11, p 77) Thus the organism is indeed an organism-as-a-whole or no organism at all

It is especially not possible to eliminate the motor side of the problem in children, as they do not have the language facility to de-

scribe their perceptual experiences, and we must therefore deal with the visual-motor pattern of drawings The child must first be able to manipulate his own hands and the tool before he can draw at all. It appears that the motor activity develops first or at least independently of the optic imagery, and that the child's first productions are purely motor phenomena or merely hand movements or scribblings without meaning, and that his first satisfaction is derived from the simple rhythmic movement. Later, he adds meaning to his own scribblings, and only subsequently starts his drawings with some preconceived concept. This purely motor stage persists even at an age when the child recognizes and reacts to many complicated sensory patterns in his world.

Goodenough has reviewed the findings from the many workers in all parts of the world, including studies of primitive children (15), and they have shown a remarkably constant order of development Rouma (19) and Burt (6) have shown that, in general, backward children draw like younger children. My own studies of defective adults (3) and African Negro children (4) show similar pictures in the development of the visuo-motor pattern analyzed from the Gestalt point of view The motor pattern must vary considerably under these conditions, especially as some of the Negro children had never used paper and pencil before The defective adults which I studied in some cases had not had paper and pencil in their hands for 10 or 15 years, if ever Thus we seem justified in assuming that motility is not the large element in the graphic productions and that the organization in the sensory field itself, as well as the conceptional interpretation, are the big factors.

Analysis of drawings of children from the Gestalt point of view has been limited to a very few studies My own analysis (3) of the copied forms of the Wertheimer (27) Gestalt patterns by defective adults and schizophrenic individuals led me to conclude from this data alone that the visuo-motor patterns arise from movement, probably vortical, and that internally organized Gestalten evolve genetically by a progressive organization in connection with the integrated intellectual functions

> "It thus appears that the more primitive sensory-motor patterns are dependent upon the principles of constant motion of a whirling or vortex type with an associated radiating directional component and with a tendency to emphasize the horizontal plane Fixed points are difficult and straight lines are

not accomplished as the shortest distance between two fixed points but as an expression of radiating tendencies Crossed lines and angulated forms present great difficulties The first evidence of expressed form was shown at the second year level as little units of whirls or loops which were perseverated most freely in the horizontal plane in a dextrad direction Some tendencies for Gestalten were seen in the third year level as rectangular forms either near each other or inside each other But some of the Gestalt principles are functions of the more highly elaborated perceptional motoric capacities and only appear at the higher intellectual levels Above three years we find the tendency to accentuate the horizontal base line, to control perseveration and to produce wavy lines instead of broken ones for the representation of straight ones and there is some effort to cross lines At the five year level there is some tendency to reduce the primitive loops to points and to make straighter lines and better recognized Gestalten At the seven year level there is a capacity for many vertical and geometrical Gestalten on the principle of internal organization and crossing of forms At the eight year level nearly all *Gestalten are possible, slanting forms are recognized but their* relation to the whole figure presents difficulties Pairing as determined by slight differences is not accomplished The major difficulties at this level seem to be in relationships of parts to wholes These things are satisfactorily accomplished above the ten year level At all levels all of the original principles are in evidence and tendencies to revert back to them are always present."

It was also found that in schizophrenia there was a dissociation due to reversions back to genetically more primitive tendencies.

I have also had the opportunity to analyze the productions of some fifty South African children examined by Dr H Nissen in connection with his work under the auspices of the Yale Psychobiological Laboratories. He used the Army performance test for memory of drawings on children in South Africa, many of whom had never used a pencil before But many of the same principles were found as in my other study Thus it was found that Gestalten seemed to arise from loops which are perseverated or concentric. Below this level there was the tendency to fill the space offered with scribblings made by large arm movements without regard to the pattern offered. Lines first occurred as horizontally continued segments of loops in a wave-like relationship Integration occurred by the combination of

these forms into related loops, angles, and crossed forms Perfection in form, correct number of parts (by inhibition of perseveration), and proper relationship of parts are added later. Slanting forms, angles other than right angles, non-rhythmic relationships, and irregular forms are most difficult

An interesting confirmation of the theory (Koffka) that development comes from maturation and not learning is brought out by comparing the productions of one of these South African Negro children of 12 years of age and a mental age of 11.5 years with the same type of production of an American Negro child of a similar chronological and mental age The productions were almost identical

Street (23) standardized a Gestalt completion test and concluded that it measured a specific capacity probably involved in perceptual processes He used pictures of different complexity in which varying parts were deleted to form the background so that, in order to perceive the picture, it is necessary to complete it by mental closure. No special principles were used in the deletions so that no definite physiological or psychological laws may be followed in completion, as, for example, those laws demonstrated by Wertheimer for proximity of parts, organization of geometrical forms, or even the organic organization of the picture itself.

In my present study I became interested in the spontaneous productions of children in the chalk drawings on the sidewalks and the open pavements of parks. The numerous small parks that dot the East Side of New York afford a great wealth of material, as, during the pleasant weather of the spring and fall, they are always covered with the drawings of the neighborhood children. Apparently no systematic study of such drawings has been made They present several advantages over the usual paper-and-pencil or slate drawings. One advantage is their absolute spontaneity. The children draw for the fun of the thing and with apparently no other goal in view than the immediate joy of activity and production except in case of the sidewalk games such as hopscotch, where the drawing affords the setting for the game

The sidewalk drawings also give us an opportunity to study drawings produced under different motor conditions. The child is usually either sitting or kneeling on the ground on which he is drawing, or leaning over from a standing position, or even precariously balancing himself on roller skates and drawing with large arm sweeps.

But the pictures produced in this way are in many ways quite similar to those described so extensively in the literature from the usual paper-pencil method

Perhaps the most important difference is dependent upon the unlimited amount of space. A child with plenty of chalk, with pavement all about him, with plenty of time, and totally unconscious of any supervision, rarely draws a complete, isolated figure He draws and draws and scribbles all over the place, delighted when his scribbles display some unexpected form, experimenting with this, he modifies it first this way and then that by some such simple variation as enlarging a loop or extending a line, often leaving some fragment of a design incomplete in order to try a new variation He may for some reason leave this place and at some subsequent time take up the game again in some new area, or other playmates may adopt his

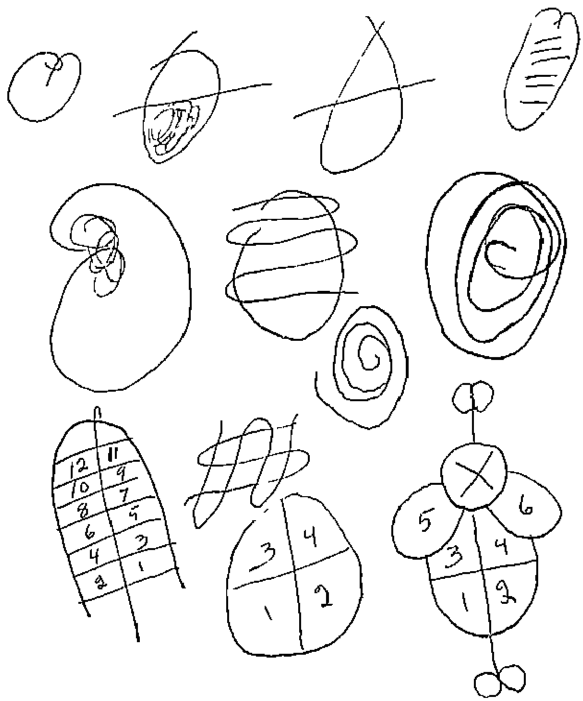


Figure 1

PLATE 1

idea and play with it in their own way. Thus definite schools of design may prevail through a neighborhood for days at a time.

It certainly becomes evident that often the satisfaction derived from these sidewalk productions is largely a motor one. This is especially true of the smaller children, but it also occurs in the older ones. They may simply make large arm movements with swaying circular lines until they have filled all the nearby space or have used up the chalk (Plate 1) At times they whirl themselves about, making large circles, and then they may fill them in block-wise At other times simple designs or directional features are shown such as making one block of cement white with the others left as background, or lines are drawn from various fixed points on the pavement, or the child simply moves along drawing lines in the direction in which he is going Numbering things is a frequent pastime possibly closely associated with the obsessional trends of children; blocks of city pavement are numbered consecutively, or simple designs as oblongs are made and marked off in parts and numbered. Usually any given area is filled with the results of many of these different activities as well as with some of the better-organized pictures of older children. In the latter case loops and circles in many variations are gradually worked into more elaborate designs. Sully (24) has emphasized that the first designs are circular and lunar forms of the human face. Shinn's (22) niece recognized *O*'s at 12 months and *C*'s and *Q*'s soon after My own studies of defective adults and Negro children showed that whirls and loops were characteristic of the lowest levels. Dora Musold (16), working under Volkelt, tested children for their ability to discriminate the size of spheres, surface circles, contour circles, and straight lines She found that the small children had a better discrimination for spheres than the adults or older children but had a much poorer discrimination for straight lines. This is further evidence to show that, in the developing mentality, spheres (in the case where three dimensions are used) or circles are biologically better perceived than straight lines. The child is first acquainted with his own and mother's body, his own feces, his mother's breast, and his mother's face Sully has made the interesting comment that the earliest pictorial forms have embryonic characteristics. We are impressed with the probability that the postural model of the body (Schilder, 20) is the first perceptual experience and helps to determine the organization of the visual field, for, as Koffka says, "it is not the simplest form but those biologically most important

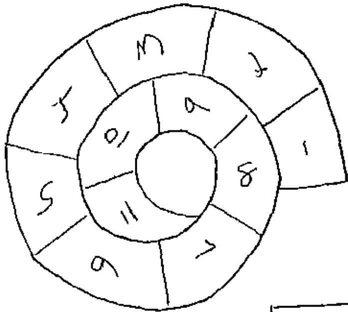


Figure 1

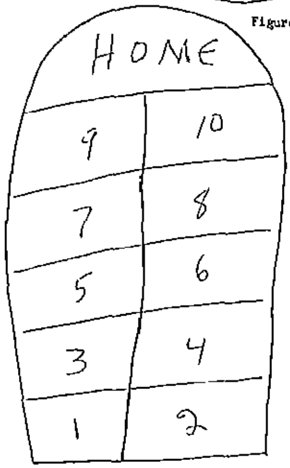


Figure 2

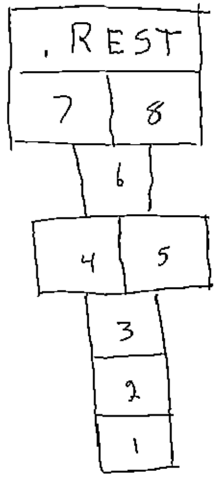


Figure 3

PLATE 2

which are first evident in infantile perception." However, the visual field may have underlying physiological features of its own that determine its organization into movement of a whirling, circular, and wave type Kammer and Schilder (10) have shown that the characteristic properties of optic imagery are movements of a wavy and circular nature, with scintillations and multiplications or fading and diffusing of the image, or parts of it, and participation of the background in the same process.

Before we go into a discussion of the more elaborate pictorial designs of children I want to discuss a group of their sidewalk games, known as hopscotch Hopscotch is a game that has several variations,

but, after watching the children play it, I am convinced that the variations are dependent upon the different age levels, that is, on the maturation level of the perceptual motor pattern Of course, this is not absolute, as older children will sometimes play with a younger group and vice versa, and older groups of children will sometimes revert to simpler forms, and younger children will sometimes emulate their elders. But, in general, small tots just able to hop on one leg will make a simple whirl design on the pavement, mark it off in blocks (Plate 2, Figure 1), and hop from block to block until they reach the center The aim is to hop always on the one foot without putting the other foot down and without touching any line The size of the loop will depend on the abilities of the members of the group, and it is readily enlarged or made smaller at the open end. This is a very simple game, but apparently it affords considerable satisfaction for tots of four years to six or seven years of age One sees in this an almost pure example of a perceptional motor game involving the principles of a visual pattern of the primitive whirl type and a motor pattern involving the simplest rhythmic hopping in a whirling direction Whirling is an activity which young children apparently enjoy very much and which is probably dependent upon their primitive postural reflexes and vestibular sensations (21). Similar whirling games were also played by Kohler's (12) apes.

The more mature children make a hopscotch design like Figure 2 of Plate 2. The general outline is more rectangular, although it is rounded at the top where "heaven" or "home" or "rest" is located. The rest is marked off by cross lines into eight or ten square sections The aim here is more complex. A block is also used, and the child hops and kicks the block from one square to the next until it reaches the top without the child's touching a line or putting down the other foot; the block must be kicked each time into the proper spaces without letting it stop on a line This involves several more complicated processes. The primitive circular form has organization only from the periphery to the center and in a clockwise or counterclockwise direction But in this second form we have a top and bottom and right and left side, and there are also straight lines crossing each other. The motor pattern requires a better adaptation and includes the accurate control of an inanimate object and also a back-and-forth direction.

A still more sophisticated form of this game is played by girls of eight or ten years. They make a design such as is seen in Figure 3,

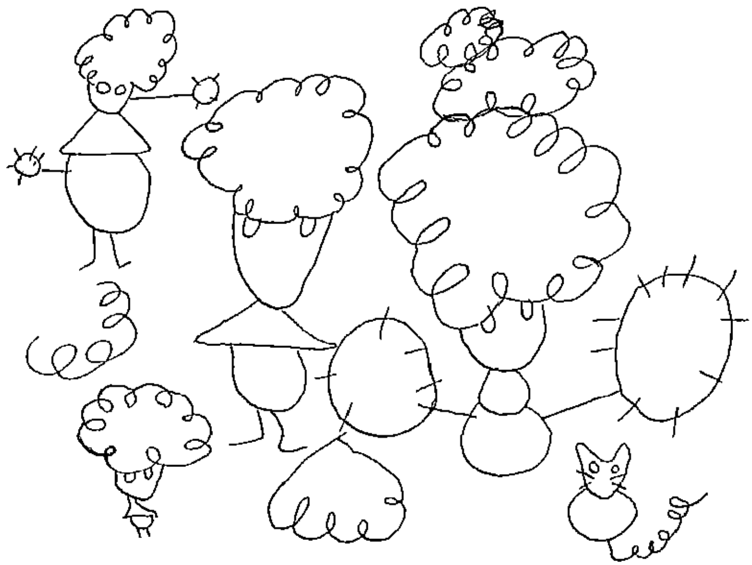


PLATE 3

Plate 2 Here the pattern includes the alternate position of form, and the motor activity includes an extra step over the paired spaces where the girls will put both feet down over the paired spaces between each hop in the individual spaces, at the same time kicking the block from one space to the next. This represents merely an elaboration of the former game by adding an alternate rhythm in both the sensory and motor sides of the game.

I must call attention to the close resemblance of the various scribblings shown in Plate 1 and these hopscotch designs as shown in Plate 2 Besides this, Figure 1, Plate 1 was found in very small proportions not far from an area where the girls were playing the game as shown in Figure 3, Plate 2. Thus is seen the tendency for the games to grow out of the primitive scribblings and for design to be influenced by the games by reverting back to the simpler forms of loops.

Of greater interest are the drawings that are scattered lavishly on the sidewalks in a spirit of experimentation such as are seen in Plates 3, 4, and 5. In Plate 3 we see a study in whirligigs, sometimes just for their own sake, as in the lower left-hand corner, or as the

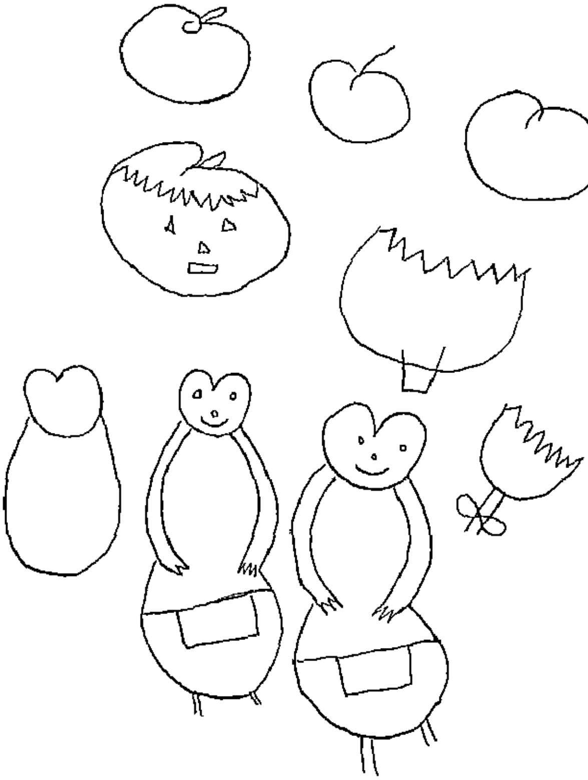


PLATE 4

tail of a cat, as in the lower right-hand corner, but more often as the central theme in a human figure, in which it is featured as the hair The rest of the body is suppressed in different ways, once the legs and feet are missing but the hands are exaggerated, again the arms are missing but the legs are better represented, but at best are unimportant as compared to the hair. And even the arms, when present, are so casually placed as to appear connected with the face or belly The body is always shown in a similar way, as is also the face, in which only the eyes appear as features Even in the cat a similar pattern is shown. In different parts of the same park, different modifications of this design were seen and continued to

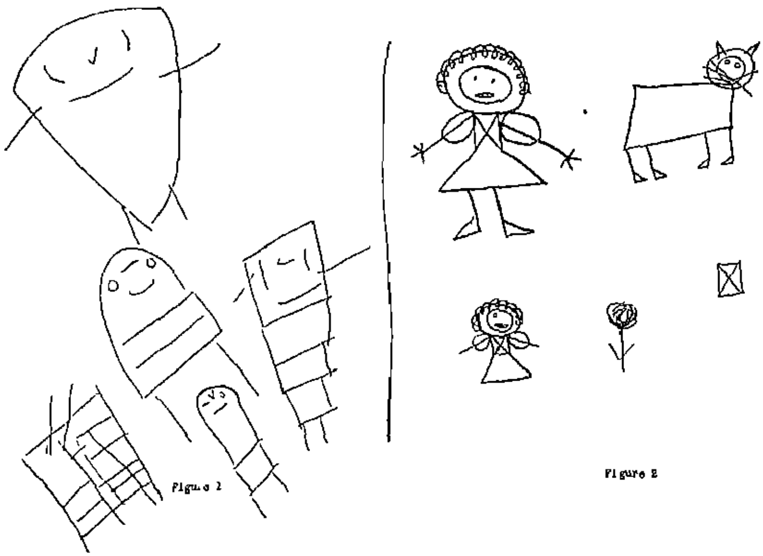


PLATE 5

appear for several days. One might look upon it as a period in which the school of whirligigs dominated the art tendencies in this particular locality

Plate 4 shows a *Halloween study. It was found at Halloween* time. The simple incomplete primitive loop is used with variations to form an apple, pumpkin, jack-o'lantern, a bouquet, and several human forms. The surrounding neighborhood was covered with many more of these in various forms of fragmentation and evolution

Plate 5 represents an experiment in angular creatures Its similarity to the second form of hopscotch is evident, although no hopscotch was seen *in this park at this time.*

A well-developed girl of about seven years was seen experimenting with triangles and squares in the study in Figure 2, Plate 5 One might comment along the way, that the modern tendencies in cubist art represent about a seven-year level in the maturation of pictorial symbolism. It is as though the cubist said "Lo and behold, I have outgrown the infantile level of loops and circles and can now accomplish and appreciate straight lines and crossed lines!"

The more sophisticated drawings of the older children are often

isolated figures and do not differ much from the paper-and-pencil drawings which have been so extensively studied and finally standardized by Goodenough. There are, however, some important points that may at least be emphasized As a rule, the child does not draw man-in-general with the interest focused on the correctness of the details, neither is it any special man with the emphasis on the form, rather is it a man doing something As Guillet (9) found, the small boy is chiefly interested in what the men and animals he sees in pictures are doing, and he is less interested in their form or any other detail. The Stanford-Binet tests show that the young child defines things in terms of their use, and, as Koffka says, to the child the man is not made up of his members but the members belong to the man, therefore, if the picture drawn is depicting some concept that does not utilize some member, that member is likely to be deleted in the child's drawing (Plate 4) Thus the man walking does not need arms, and the man urinating needs not only genitals but also the umbilicus which is suggested by these, but which is usually not shown in the drawings of men; and the man with a hat does not need legs, etc. Nevertheless, one is impressed with the pictorial success of the prominent idea, as in the case of the urinating man—with his spread legs and even the expression on his face. The man walking gives us an excellent example of the influence of the sensory field on the pictorial forms It has been experimentally shown by Kanner and Schilder (10) and Wertheimer (26) that the perception of motion is accomplished in the optic imagery by frequently recurring (or many) static figures Or as the 27-months-old child, Ruth, said when remonstrated with by her aunt for putting so many little dashes for legs about the round loop that she called a mouse, "I give dat mouse many legs so he can run wight away" (Brown, 5). That children's drawings are sometimes apparently symbols of unconscious concepts is shown in Figure 3, Plate 6, which represents a fragment of a very elaborate design made up of many interwoven forms like the simplest circular hopscotch, with the cross lines apparently representing many stairs up which girls were climbing at more or less regular intervals That they were girls is evident by the breasts and umbilicus (infantile vagina?) and the curls There were about 25 such girls, all obviously similar, though varying by lacking in different details; some without hair or arms or legs but all showed the sex features At the top of this intricate

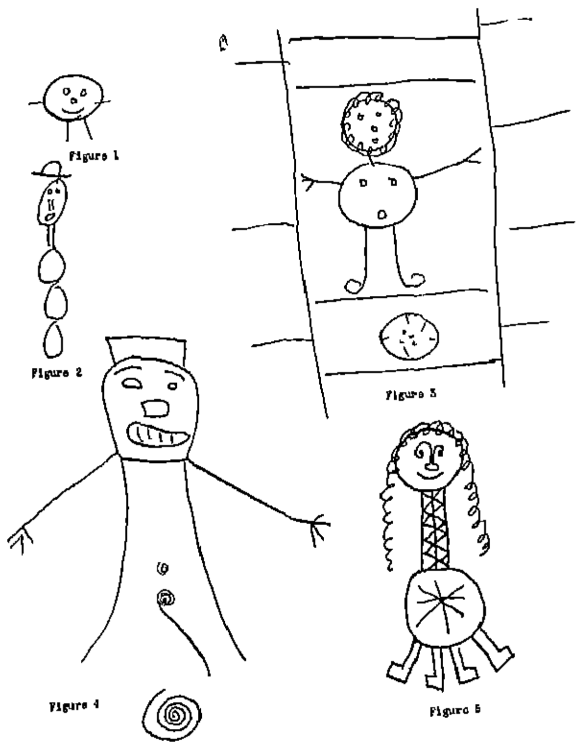


PLATE 6

stair design (inside of the many loops) was a house. That this is a sex fantasy is evident.

In psychological terms we may say that the goal in children's drawings is to establish an equilibrium between the mental symbols as determined by the biological background of the perceptional motor pattern at the different stages of the maturation of the organism as a whole and the reality of the world as it is perceived The traces that are used must correspond not alone to the stimuli that produce them but also to the organization characteristics of the sensory field. Wulf and Koffka (28) have done some experiments on traces that bear on this subject. They showed a variety of simple geometrical forms to adults and asked them to reproduce them after intervals of from 24 hours to a month or more. In many successive productions

there were constant progressive changes towards a simplification or exaggeration of the figure but generally towards a balanced or symmetrical form It thus appears that the traces that were originally produced by the stimuli were modified by the biological character of the field in favor of a closer organization. Koffka (12) has said, "The development of such configurations can not be conceived as a simple combination of sensations or as an outward manifestation of a juxtaposition of repeated sensations but as a result of a process which alters, refines, recenters, enriches the configuration throughout its entire make-up—a procedure in which maturation participates largely," and, we may add, organization also participates as a function of the sensory field as emphasized by Kohler, probably also the motor pattern, the postural model of the body, and other sensory experiences may participate

References

1 AMENT, W The mind of the child a comparative history of life. London: Simpkin, 1926 Pp 78

2 BALDWIN, J M. Mental development. New York Macmillan, 1903 Pp. xvi+496

3. BENDER, L Defective adults *Arch Neur & Psychiat*, 1932, 27, 661-686

4. ——— African Negro children. (In preparation.)

5 BROWN, E. E Notes on children's drawings *Univ Calif Stud.*, 1897, 2, Pt 1 Pp 75

6 BURT, C Mental and scholastic tests. London London County Council, 1921. Pp vii+239

7 CUSHMAN, L S The art impulse, its form and relation to mental development *Nat. Educ Asso*, 1908, 515

8. GOODENOUGH, F L Measurement of intelligence by drawing Yonkers, N Y. World Book Co, 1926 Pp xi+177

9. GUILLET, C Retentiveness in children and adults. *Amer. J. Psychol*, 1909, 20, 318-352

10. KANNER, L, & SCHILDER, P Movement in optic image *J Nerv & Ment. Dis*, 1930, 72, 489-517

11. KATZ, D Der Aufbau der Tastwelt *Zsch f Psychol*, 1925, Ergbd 11. Pp xii+270

12 KOHLER, W The mentality of apes (Trans by E Winter.) London: Kegan Paul, 1924 Pp viii+342

13 ———. Gestalt psychology New York. Liveright, 1929 Pp x+403

14 KOFFKA, K The growth of the mind. (Trans by R M Ogden) New York Harcourt, Brace; London Kegan Paul, 1924 Pp xvi+383

15 LAMPRECHT, K Les dessins d'enfants comme source historique *Bull. de l'Acad. roy de Belgique*, 1906, Nos 9-10, 457-569.

16. Musold, D, Buhler, K, & Volkelt, H. Die neue Psychologie *Zsch f Psychol.*, 1926, 99, 145-154.

17 Paulsson, G The creative element in art *Scand Sci Rev.*, 1923, 2, 11-173

18. Preyer, W The mind of the child (2 vols) New York Appleton, 1888, 1889. Pp 346, 317

19 Rouma, G. Le langage graphique de l'enfant Bruxelles. Misch & Thron, 1913 Pp 283

20 Schilder, P Der Korperschema Berlin Springer, 1923 Pp 92

21. ——— Brain and personality (*Nerv & Ment Dis Monog Ser.*, No 53) New York & Washington Nerv. & Ment Dis Publ Co., 1930 Pp 130

22 Shinn, M W Notes on the development of a child *Univ Calif. Stud*, 1893-1897, 1, Pts 1-4 Pp 178

23 Street, R F A gestalt completion test. *Teach Coll Contrib Educ.*, 1931, No 481 Pp vii+65

24 Sully, J Studies of childhood New York Appleton, 1897 Pp 527.

25 Verworn, M Kinderkunst und Urgeschichte. *Korresponden d dtsch anthrop Gesellsch.*, 1907, 42-46

26 Wertheimer, M Experimentelle Studien uber das Sehen von Bewegung. *Zsch f Psychol*, 1912, 61, 161-265

27 ——— Untersuchungen zur Lehre von der Gestalt *Psychol Forsch.*, 1923, 4, 301-350

28. Wulf, F., & Koffka, K Ueber die Veranderung der Vorstellungen *Psychol Forsch*, 1922, 1, 333-337

Bellevue Hospital
New York City

LES PRINCIPES DE LA GESTALT DANS LES DESSINS SUR LE TROTTOIR PAR LES ENFANTS

(Résume)

Les dessins spontanes sur les trottoirs et les pavages par les enfants et de tels jeux que la marelle offrent un excellent matériel pour l'analyse psychologique de la genèse et de la maturation de la perception de la forme et de son expression motrice chez les enfants Chez les jeunes enfants les dessins et les jeux sont principalement une forme motrice de l'expression des balancements rythmiques des bras et des rotations rapides du corps Les formes résultantes qui ressemblent aux boucles et aux rotations acquièrent peu à peu une signification pour l'enfant Bien des experiences spontanées avec diverses déviations et combinaisons des formes les plus simples sont typiques de l'évolution aux dessins moins naifs des enfants plus âges. Ainsi ce qu'éprouve l'enfant est basé sur la récognition perceptive et donc conceptive de certains groupes de formes ou Gestalten qui lui offrent une signification Ceci n'est pas base sur son analyse ou sur son évaluation des détails tels quels mais sur leur relation les uns aux autres et à son monde entier Ses dessins représentent son habileté à

faire des symboles qui représentent ces relations des parties et leur signification totale pour lui. Le but de ses dessins c'est d'établir un équilibre entre ces symboles mentaux comme déterminé par le fond de la forme perceptive motrice aux divers stages de la maturation de l'organisme total et par la réalité du monde perçu par lui Les traces employées doivent correspondre non seulement aux stimuli qui les produisent mais aussi à l'organisation caractéristique du champ des sens et sont aussi influencées probablement par la forme motrice, le modèle de posture du corps, et d'autres expériences sensorielles

BENDER

GESTALTPRINZIPIEN BEI DEN ZEICHNUNGEN VON KINDERN AUF DEM TROTTOIR

(Referat)

Die spontanen Kreidezeichnungen von Kindern auf Trottoir und Pflaster, und Kinderspiele wie z B. Linienuberspringen (hopscotch) liefern ausgezeichnetes Material zur psychologischen Analyse des Anfangs und der Reifung (maturation) der Perzeption der Form und ihre motorische Äuserung bei Kindern Bei jungen Kindern sind sowohl das Kritzeln wie die Spiele eine motorische Ausserung des rhythmischen Schwingens der Arme und des Wirbelns des Korpers Die resultierende Schling- und Wirbelartigen Zeichnung nehmen allmahlich fur das Kind Bedeutung an Sehr viel spontanes Experimentieren mit verschiedenen Variationen und Verbindungen der einfacheren Formen ist fur die Entwicklung zu den komplizierteren Zeichnungen der alteren Kinder typisch Die Wahrnehmungen des Kindes stutzen sich also auf die Erkennung, durch Empfindung und daher durch Auffassung, gewisser Konstellationen oder Gestalten die fur es Bedeutung haben Letztere grundet sich nicht auf des Kindes Analyse oder Verwertung der Einzelheiten als solche, sodern auf ihre Beziehung zu einander und zu seiner ganzen Welt. Seine Zeichnung ist eine Darstellung seiner Fähigkeit, Symbole zu zeichnen, die diese Beziehungen der Teile und ihre gesammte Bedeutung fur ihn darstellen Der ziel seiner Zeichnungen ist die Errichtung eines Gleichgewichts zwischen diesen geistigen Symbolen, in ihrer Bestimmung durch den Hintergrund der perzeptiven-motorischen Gestalt in den verschiedenen Stadien der Reifung des Organismus als Ganzes, und der Wirklichkeit der Welt so wie sie wahrgenommen wird Die verwendeten Linien (traces) mussen nicht nur den sie verursachenden Reizen, sondern auch den Organizationseigenschaften (organization characteristics) des Sensoriums (sensory field) entsprechen, und werden wahrscheinlich auch durch die motorische Gestalt (pattern), das Haltungsmodell (postural model) des Korpers, und andere sensorielle Wahrnehmungen beeinflusst

BENDER

SOCIAL FACILITATION OF FEEDING IN THE ALBINO RAT*

From the Psychological Laboratories of Stanford University and the University of Wisconsin

H F Harlow

The effect of social influences on the drives of the normal human being is quite obvious. We inhibit certain tendencies and reinforce others in the presence of individuals like ourselvves Eating is influenced, probably not so much as to quantity as to appreciation. A good meal tastes better if we eat it in the company of friends. Likewise, it tastes much worse if we eat it in disagreeable company.

In the study of animals we are reduced to a more quantitative attack on the problem. It is impossible to determine the rat's increased appreciation of food (if any such exists) but we can measure the amount ingested It has seemed desirable to study the effect of a social situation on this elementary type of response by comparing the amounts of food taken in solitary feeding as compared with the amount eaten by the same animal when feeding in a group.

Pertinent Literature

Fischel (2) was able to show that hens, though exhibiting no signs of hunger, were stimulated to begin eating again by the sight of another hen feeding Bayer (1) completed a series of researches designed to check and extend this observation In these investigations,

> "The research was carried out as follows a hungry V-animal (experimental animal) was placed before a large heap of wheat (a food well liked by the hens) and allowed to eat until satiated Then a hungry A-animal (exciting animal) was introduced "

Two measures of the effect of the social facilitation of the feeding responses which were subsequently obtained were made by Bayer The first of these was the changed behavior of the hen elicited by the above situation, and the second was the amount of food that the V-animal ate after the A-animal was admitted to the cage.

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, January 19, 1932

Since one of the two hens almost invariably dominated the other, Bayer arranged the situation so that the dominant animal was alternately the V-animal and the A-animal.

The behavior of the hens is described by Bayer as follows:

> "As soon as the A-animal was placed in front of the food, it began to eat with great zeal. Now how did the V-animal behave? If it were the dominant animal it began to attack the A-animal immediately. To one who is not a behaviorist this obviously indicates that the V-animal attempts to prevent the A-animal from eating as a result of envy over the food (*Futterneid*). As soon as the V-animal noticed, however, that it was having little success, and that the A-animal continued to eat in spite of everything, it began once more to eat the food, intermittently striking at the other animal. This behavior was facilitated by the fact that the V-animal no longer ate with its original zeal. If the V-animal was the subordinate animal, it made no effort to hinder the feeding of the other upon its entrance. Instead, it began to eat again in spite of the fact that the dominant animal would strike at it from time to time. The sight of the dominant animal eating was apparently a strong enough stimulus to inhibit the fear which the subordinate animal normally felt. One could note that the subordinate animal behaved cautiously and ate, so to speak, 'behind the back' of the other."

The social increment as measured by the increased amount of food eaten depended somewhat upon the particular situation. Using one V-animal and one A-animal, Baer experimented on eight hens. Social increments (amount of food eaten by the V-animal after the A-animal was admitted to the cage) were obtained in every case and averaged 34%, the lowest being 25%, and the highest 43% of the amount eaten to induce satiation.

Experiments on the same animals in which the situation was altered so that there was always one V-animal and three A-animals gave even greater social increments, averaging 53%, and ranging from 33 to 67%.

Where the situation was reversed so that there was only one A-animal and three V-animals, the effect was very much less. One out of four hens was entirely passive, and no animal showed a social increment in excess of 33%. With the three V-animals eating in the same compartment, results obtained on four animals gave a social increment averaging 10%, and, with the three V-animals

eating in separate compartments, results obtained on four animals gave a social increment of 21%.

A fourth experimental situation was so arranged that four hens would eat separately one day and together the next. Two experiments were completed and social increments (increased consumption of food for day 1 over day 2) were obtained in every instance. These ranged in the individual animals from 33 to 200%, and averaged 96%

Seven experiments have been conducted.[1] In the following study, all have the same purpose but vary slightly in technique. For purposes of convenient understanding, the technique and results will be given for each separately before any attempt is made to compare the results or draw general conclusions

Technique and Results

Experiment 1 In this experiment the animal was allowed to feed to satiation, an hour period being assumed to meet this requirement. At the end of this time a hungry rat (24-hour food deprivation) was introduced and allowed to feed for an hour. The object was to find if the satiated rat would be stimulated to feed further in the second hour by the presence of the hungry rat Corn was used as food which made it easy to observe the amount eaten by the satiated animal as the grains seized could be counted The results were tabulated for eight animals on each of three days, and in no case were they distinctly favorable to the existence of a social facilitation process, as the satiated animal ate no more in the second hour in the presence of the hungry rat than when alone.

Experiment 2. In this experiment young rats without previous experience with solid food were used Tweny rats were weaned at the age of 18 days, and, as the mothers had been fed outside the cage for the three preceding days, they could have had no experience with solid food They were fed a liquid diet for a certain number of days (2, 7, 12, and 17, respectively, for the four groups of five rats each) so that, on the day of starting the experiment, the groups were 20, 25, 30, and 35 days old. They were then placed in individual cages and fed corn during two one-hour periods per day

[1]Experiments 1-6 were conducted at Stanford University under Dr C P Stone The work was financed by the Thomas Welton Stanford Fund Experiment 7 was done at the University of Wisconsin

TABLE 1
COMPARISON OF AMOUNT OF CORN EATEN AND SPILLED BY ALBINO RATS FEEDING SEPARATELY AND IN GROUPS OF FIVE

Group No.	Age days	Individual feeding Grains eaten	Individual feeding Grains spilled	Group feeding Grains eaten	Group feeding Grains spilled
1	24-25	14	67	21	204
2	29-30	18	59	30	221
3	34-35	21	84	41	242
4	39-40	24	78	49	237
Total		87	288	141	904

for five days. On the sixth day, all five animals were placed in one cage and allowed to feed.

Objective comparison was made between the fifth day (last individual feeding) and sixth day (first group feeding) The criteria used were the number of grains eaten and the number of grains spilled on the floor The results are summarized in Table 1.

The number of grains eaten by the 20 rats on the last day of individual feeding was 87 On the first day of group feeding the total number of grains eaten jumped to 141. This amounts to an increase of about 70%. The amount of corn spilled was even more strikingly increased. The 20 rats spilled 288 grains when feeding individually, but when feeding in groups they spilled 904 grains, an increase of about 200%

The behavior of the rats was even more markedly different. The rat eating alone characteristically fed leisurely and with occasional wandering about the cage. In the group situation, the performance was hurried and apparently competitive. Each rat would hasten to a corner, or two rats would eat back to back The process of getting grains from the pan occasioned much scrambling and pushing, which was the chief cause of the large amount spilled The group performance appeared much more highly motivated than the process of individual feeding.

Experiment 3 In this experiment 20 animals, all males, were reared in isolation from the age of 20 to 40 days. From the 40th to the 44th day, inclusive, the amount of standard diet eaten by each animal in his two feeding periods was recorded, and on the 44th day they were grouped by weight into ten pairs. From the 45th to the 49th day they were fed in pairs and the total amount

eaten by the pair compared with the sum of the amounts eaten by the two individually Some of these pairs showed a social increment, some no effect, and some a social decrement. An observation of the behavior of these animals indicated that in many cases the two would spend the feeding hour in play, even though a steady decline in weight indicated that they must be in a state of hunger.

Experiment 4. In this experiment 20 males, 40 to 45 days old (with two exceptions, 61 days), were used The situation was the same as in the third experiment, with the chief exception that previous to the beginning of the individual feeding, they had been housed and fed in groups of four or five, so that they were accustomed to a social feeding situation

The experiment lasted 25 days, including five individual days, five social, five individual, five social, and a final five individual days. These will be designated control 1, experimental 1, control 2, etc

The results on amount of food eaten are summarized in Table 2 The differences here are strikingly in favor of the concept of social facilitation The mean difference between the first control and the first experimental period is 13 35 grams and the *S D* of this difference is only 1 69, giving a critical ratio of 7 90. The other differences are even greater, and the critical ratios correspondingly higher.

The amount of gain in weight made by each rat was also computed These results are given in Table 3. The differences here are all large and reliable except for the comparison of experimental 1 and control 1. The conclusion which is obvious from these two criteria is that social facilitation of feeding responses unquestionably does occur, and that it is important

Experiment 5. This experiment differs from the previous one in that only nine animals were used, and during the two experimental periods they were fed in groups of three Since the number of cases is so small, the results are not given, but the critical ratios were significant for both food eaten and weight gained, in every comparison of the group and individual situations The results did not indicate, however, that there was greater facilitation when three animals were feeding together than in Experiment 4 in which they fed in pairs.

Experiment 6. Since Bayer (1) claimed that "envy" was one of the factors operating in the social facilitation process, an attempt was made to eliminate the competitive factor A cage, 7 inches

TABLE 2

AMOUNT OF STANDARD DIET EATEN BY YOUNG ALBINO RATS FEEDING SEPARATELY AND IN GROUPS OF TWO

Period	Total grams eaten	Average grams eaten	
Control 1	917	45.85	
Experimental 1	1184	59 20	
Control 2	794	39 72	
Experimental 2	1220	61 10	
Control 3	797	39 85	
Mean differences	Diff.	*S D* *diff*	*C R*
E1 — C1	13 35	1 69	7 90
E1 — C2	19 50	1 67	11 67
E2 — C2	21 30	1 46	14 58
E2 — C3	21 15	1 42	14 89

TABLE 3

AMOUNT OF WEIGHT GAINED BY YOUNG ALBINO RATS FEEDING SEPARATELY AND IN GROUPS OF TWO

Period	Total grams gained	Average grams gained	
Control 1	294	14 70	
Experimental 1	333	16 65	
Control 2	50	2 50	
Experimental 2	459	22 95	
Control 3	75	3 75	
Mean differences.	Diff	*S D* *diff*	*C R*
E1 — C1	1 95	1 41	1 38
E1 — C2	14.15	1 54	9 18
E2 — C2	20 45	1 93	10 59
E2 — C3	19.20	1 48	12 97

long, 3 inches high, and 1½ inches wide, was used to confine a large adult rat An opening in the front end allowed him to reach out his head, but he could not escape or turn around When, thus confined, he was placed in the cage with a feeding rat, the other rat was not stimulated to eat more, whether or not the confined rat was placed near enough to the food to eat from his position It does not seem likely, then, that "envy" or personal animosity is a factor operative in the social facilitation situation.

Experiment 7 The first six experiments were considered as more or less exploratory, and an attempt was made to duplicate Experiment 4 under more carefully controlled conditions Since we consider this the crucial experiment of the series, the technique is detailed fully

The rats were weaned at 18 days and reared in individual cages until 30 days old Their two daily feedings were one-half hour in length, at 9 A.M. and 5 P.M A double pan was used for feeding, with food only in the inner pan, so that most of the spilled food was collected in the outer. Food falling on the floor of the cage was caught by a sheet of oilcloth placed underneath.

Only the amount of food eaten was taken as a measure of facilitation. Exactly 20 grams of food were furnished the animal daily, and the amount eaten was computed by collecting carefully all spillage and reweighing The computations were made to 0 1 gram on scales which are accurate to 0 05 gram

At the beginning of the experimental period, 34 animals were divided into 17 pairs on the basis of weight On the 31st day of life, they were given their first group feeding, in pairs, and from then on to the 49th day, every odd-numbered day was an experimental period and every even-numbered day (32 to 50) a control day To avoid any effect of familiarity with surroundings, the animals were alternated between the home cage of one member of the pair and the other

Table 4 summarizes the results of this experiment. In the social situation the average amount of food eaten was 13 59 grams, while in the individual situation only 12.53 grams were eaten The difference is 1 058, and is 3.59 times its standard error. It is therefore a statistically "true" difference

It seems important not only that the group means are significantly differentiated, but also that the difference from day to day is consistent The differences and critical ratios for each two-day sequence have been computed, but in general (as a result of the small number of cases, no doubt) these ratios are less than three A definite tendency is shown, however, for group feeding to manifest a significant facilitating influence upon the amount of food ingested

The behavior of the animals in this experiment (alternating daily between group and solitary feeding) manifested even more excitement than in the previous experiments in which five-day periods had been used.

Discussion

The results of the several experiments argue clearly for a process of social facilitation of the feeding response in the rat. If we rely upon differences in food ingested, in food spilled or in weight gained, we see that there is in practically every case a significant superiority of the group situation.

The observations of behavior also indicate a process of facilitation The animals in the group situation were characteristically highly motivated, displaying an excess of activity in getting food, struggling, crowding other rats from the food dish, etc. There was a distinct difference in the leisurely feeding of the individual rat, which was frequently interrupted by exploration of the cage, etc.

These generalizations hold true only when the various rats are

TABLE 4

Average Amount Eaten by Pairs of Albino Rats Feeding Together and Separately on Alternate Days

Day of experiment	Group feeding	Individual feeding
1	10.35	
2		9 22
3	11 67	
4		10 43
5	12.57	
6		11 20
7	12.74	
8		12 09
9	13 24	
10		12.24
11	14.08	
12		13 32
13	13 66	
14		13.56
15	15 42	
16		13 83
17	15 84	
18		14 69
19	16 33	
20		14 75
Average	13 59	12 53
Difference	1 057	
$S.D_{diff}$	0 294	
C R	3.59	

free and competing. The experiments in which this condition did not prevail (Experiments 1 and 6) do not show social facilitation.

The observation of the behavior of these animals in the individual and social situations leads to another conclusion which is of considerable importance in the understanding of hunger motivation. It has often been believed that the immediate stimulus for ingestion of food was the stimulation derived from contractions of the stomach. Yet it is well known that struggling and excessive striped muscle activity tend to inhibit the contractions of the smooth musculature. We have, then, in this experiment, the anomalous condition of animals in which the stomach contractions have undoubtedly been inhibited or diminished eating greater quantities of food than they did when under comparable conditions without the external disturbance. *We wish to suggest, then, that ingestion of food is* determined more by the external situation than by the actual interoceptive stimulation. This conclusion is substantiated by much common knowledge of human behavior, such as the influence of music, pleasant surroundings, the holiday atmosphere, etc

Basic Mechanisms of Social Facilitation

1. *Social facilitation does not depend upon learning.* Rats without previous experience with solid food (Experiment 2) and rats without previous experience with food in the presence of other rats (Experiment 7) manifest the effects of facilitation on the first day of the experiment

2. *The process is not subject to change and adaptation.* Five periods of five days each (Experiment 4) and 20 days of alternate feeding (Experiment 7) showed differences at the end of the experiment equal to or greater than those at the beginning

3. *It does not depend upon imitation* A satiated rat watching a hungry rat eating (Experiment 1) is not thereby motivated to eat more.

4. *It does not depend upon "envy"* The presence of a restrained, non-competing rat (Experiment 6) does not motivate the free rat to ingest larger amounts of food

5 *It is not a function of the size of the group.* Experiment 2 (five rats in group), Experiment 4 (two rats), and Experiment 5 (three rats) show differences of similar size

6. *It may be related to age.* Table 1 (Experiment 2) indicates that the older rats increase their food intake proportionately more

than the younger rats Evidence from the individual records of some of the other experiments, however, does not substantiate this finding.

7 *The essential condition for the occurrence of social facilitation is the presence of rats unrestrained and actively competing with each other for food.*

Summary

1. The effect of social facilitation upon the feeding response in the albino rat has been demonstrated.

2. This social facilitation takes place only between rats that are unrestrained and freely competing

3 Facilitation is independent of previous experience

4. It does not depend upon imitation or envy.

References

1 Bayer, E Beitrage zur Zweikomponententheorie des Hungers. *Zsch. f. Psychol*, 1929, **112**, 1-53

2 Fischel, W. Beitrage zur Soziologie des Haushuhns *Biol Centbl*, 1927, **47**, 678-696

University of Wisconsin
Madison, Wisconsin

LA FACILITATION SOCIALE DE L'ALIMENTATION CHEZ LE RAT BLANC

(Résumé)

Sept expériences faites sur le comportement alimentaire des rats, qui ont mangé individuellement et en groupes, montrent de l'évidence constante de l'existence de la facilitation sociale, si on la mesure par la plus grande quantité de nourriture ingerée, la quantité de nourriture versée, ou la quantité de poids gagné. Les différences ont été statistiquement constantes dans la plupart des expériences.

Les expériences individuelles indiquent que la facilitation sociale des réponses alimentaires ne dépend pas de l'expérience antérieure, qu'elle n'est sujette ni au changement ni à l'adaptation (en une durée de 20 jours), et qu'elle est relativement indépendante de la grandeur du groupe compétiteur La condition essentielle pour la facilitation sociale semble être la présence d'un rat compétiteur non restreint

La facilitation sociale ne dépend pas de "l'envie," comme Baer a suggéré comme résultat de ses expériences sur les poulets, parce qu'un rat restreint non compétiteur ne fait pas que le rat libre ingère une plus grande quantité de nourriture, et parce qu'un rat affamé qui mange ne fait pas qu'un rat rassasié mange plus

L'expérience suggère que l'ingestion de la nourriture peut être relative-

ment indépendante de la faim (comme incitation interne causée par les contractions stomacales) puisque l'activité motrice luttant et excessive montrée par les animaux dans cette situation devrait inhiber les contractions stomacales.

HARLOW

DIE SOZIALE ERLEICHTERUNG DES FÜTTERNS BEI ALBINO-RATTEN

(Referat)

Sieben Experimente über das Betragen von Ratten während sie einzeln und in Gruppen frassen, erstatten übereinstimmend Beweis für das Bestehen einer sozialen Erleichterung (Bahnung?) (facilitation), mass man diese an der Zunahme in der verzehrten Nahrung, an der Quantität des verschütteten Futters, oder an der körperlichen Gewichtzunahme. In der Mehrzahl der Versuche waren die Unterschiede statistisch zuverlässig.

Die einzelnen Versuche weisen darauf hin, dass die soziale Erleichterung der Fressreaktionen nicht von vorhergehender Erfahrung abhängig ist, sich innerhalb 20 Tage nicht verändert oder anpasst, und von der Grösse der konkurrierenden Gruppe relativ unabhängig ist. Die wesentliche Bedingung zur sozialen Erleichterung scheint in der Gegenwart einer unbeschränkten konkurrierenden Ratte zu liegen.

Die soziale Erleichterung ist nicht von dem "Neid" abhängig, wie Baer in seiner Arbeit mit Hühnchen vorgeschlagen hat, den eine beschränkte, nicht-konkurrierende Ratte beeinflusst die freie Ratte nicht, grössere Quantitäten Futter einzunehmen, und eine gesättigte Ratte die eine hungrige Ratte beim Fressen beobachtet, wird hierdurch nicht beeinflusst, mehr zu fressen.

Das Experiment weist darauf hin, dass die Einnahme von Futter vom Hunger relativ unabhängig sein kann (betrachtet man den Hunger als einen inneren, durch Magenkontraktionen verursachten Trieb), da die von den Tieren in dieser Lage offenbarten Sträubungen und übermässige Geschäftigkeit die Magenkontraktionen hemmen sollten.

HARLOW

SHORT ARTICLES AND NOTES

EDITED BY

JOHN E. ANDERSON
CHARLOTTE BUHLER
CYRIL BURT
TOSHIO NOGAMI
IVAN P PAVLOV
H. PIÉRON
CALVIN P STONE

SOME FACTORS INFLUENCING THE AFTERNOON SLEEP OF YOUNG CHILDREN

RUTH STAPLES

This study was undertaken for the purpose of determining to what extent, if any, the afternoon sleep of young children may be influenced by the amount of time spent in outdoor play during the morning, by the particular person in charge of the children during the afternoon nap period, and by a short rest before luncheon The subjects were 30 children attending the University of Nebraska Nursery School, 14 girls and 16 boys, ranging in age from 2 years 9 months to 4 years 11 months

Morning Outdoor Play and Afternoon Sleep. For a period of eight months daily records were kept by the teachers for *each child*, as follows. (1) the number of minutes spent out of doors on the playground during the morning, (2) the minutes of sleep during the afternoon nap period, and (3) the number of minutes elapsing between the time of going to bed and going to sleep

In mild sunny weather the children spent the entire morning play period outdoors on the playground During the majority of school days throughout all seasons of the year a part of the morning was spent indoors and a part outside, the length of the outdoor period varying according to the weather, the temperature, whether the child had completed a project and was eager to go out, or whether he was absorbed in play and desired to finish it If the day was extremely cold or stormy the children remained inside all day The time spent outdoors varied from 0 to 2½ hours, this maximum and minimum remaining the same during all seasons of the year The playground was large, permitting a great deal of running about, and was equipped with types of play apparatus stimulating active exercise. The indoor room provided much less space, and the equipment was designed for

less active play. The children went to their cots in the sleeping room directly after luncheon (from 12 30 to 1 00) They were ordinarily allowed to sleep as long as they wished, although occasionally a child was awakened at 3:20 P.M to be ready for school dismissal

In order to determine whether or not any relationship was present between the amount of outdoor play and the length of the afternoon sleep of these children, the following procedure was undertaken The daily records were arranged into three groups, (1) those for December, January, February, March, (2) for April and May, (3) for October and November [1] Pearsonian coefficients of correlation were then worked out for each child individually, for each season, between (*a*) the minutes spent in outdoor play and the length of the afternoon sleep and (*b*) the length of the outdoor play and the number of minutes required to go to sleep This seasonal grouping was used in order to offset the possible influence of chronological age upon the length of sleep The observations were begun in the winter and continued through the spring and following fall Since the amount of afternoon sleep has been shown to decrease with increase in chronological age (1), it is probable that a child would sleep less during the spring than during the winter, and less during the fall than during the spring Since there were a greater number of long outdoor play periods during the spring and fall than during the winter, on account of warmer weather, a spurious relationship might have been indicated had the records not been subdivided into seasons as described above The seasonal grouping also tended to offset any possible variations in weather conditions which might influence afternoon sleep. Coefficients of correlation were obtained for each child, individually, rather than taking the average for the entire group of children, because of the large individual differences existing in regard to both length of sleep and time for going to sleep, and *because of variation in the number of days of nursery-school attendance* The range in the average length of nap for the different subjects was from 38 7 minutes to 104 2 minutes, and for the length of time required to go to sleep from 14 to 48 9 minutes. Records were available for 30 different children, 20 subjects during the winter, 15 during the spring, and 18 during the fall. There were available 1732 daily records for sleep and for time spent out of doors, and 1132 for length of time required to go to sleep and time spent out of doors The average number of daily records for each subject for length of nap and outdoor play was 37 1 (winter), 30 (spring), and 29.4 (fall) The average number of records for each child including both time for going to sleep and time out of doors was 24 1 (winter), 29 0 (spring), and 29.8 (fall).

---

[1]During the winter months the attendance was lower and there were more holidays than during the other seasons Consequently, the longer period was selected in order to furnish an adequate number of records for each child

TABLE 1

AVERAGES OF COEFFICIENTS OF CORRELATION BETWEEN LENGTH OF TIME OUT OF DOORS, AND (*a*) LENGTH OF NAP, AND (*b*) TIME REQUIRED TO GO TO SLEEP

	Winter		Spring		Fall	
	Time out-doors and length of nap	Time out-doors and time to go to sleep	Time out-doors and length of nap	Time out-doors and time to go to sleep	Time out-doors and length of nap	Time out-doors and time to go to sleep
Av *r*	—06	+06	—11	+.05	+02	—09
S D	.15	13	18	22	21	.19

Results Table 1 presents the averages of the coefficients of correlation and the standard deviations of these averages.

The coefficients of correlation are almost without exception all low Between the length of nap and time out of doors 23 are positive and 30 negative and vary from —4 and +26 in the winter; —43 to +13 in the spring, and from —.24 to +71 in the fall.[3] For the time required to go to sleep and time out of doors the range is from —15 to +33 (winter), —29 to +35 (spring); and —54 to +36 in the fall Twenty-one are positive, 17 are negative, and 1 is 0 While on certain occasions a long nap was taken on the same day as a long play period out of doors, the reverse occurred as frequently Of the 6 children for whom records were available for all three seasons, only 2 are consistently positive or negative throughout the year Of the 11 for whom figures were obtained for two seasons, 7 are similar and 4 are unlike There is even less consistency in the case of the time required to go to sleep and outdoor play. No differences were evident when the cases were grouped for age or for sex.

Sherman (2), using a somewhat different technique for studying this same relationship, reports that the degree of outdoor activity does not materially influence the amount or the character of the afternoon sleep of young children although he finds that they fall asleep more quickly when they play indoors all the morning than when they go out of doors In order to compare the subjects of this study with Sherman's subjects, the time for going to sleep and the length of sleep were averaged separately for the days when the children played indoors all of the morning and for the days when they spent a part or all of the time outside The results are presented in Table 2.

All of the differences are slight and do not show statistical reliability Since these subjects were older than Sherman's, possible differences in the character of the children's activity due to age may account for the slight difference in the results of the two studies in regard to the influence of indoor play on the time required for falling asleep

[3]For case E H $r = +71$ The next highest for the same season was +23

TABLE 2

	Days outdoors	Days indoors	Diff	σ_{diff}	$\frac{D}{\sigma_{diff}}$
Mean length of nap	77.4	74.5	2.9	1.6	1.8
Mean time required to go to sleep	29.9	29.3	6	9	.7

TABLE 3

INFLUENCE OF THE ADULT IN CHARGE ON THE MEAN LENGTH OF SLEEP, THE LENGTH OF TIME REQUIRED TO GO TO SLEEP, AND THE PERCENTAGE OF NAPS MISSED

	Teacher X	Teacher Y	Difference	σ_{diff}	$\frac{D}{\sigma_{diff}}$
Mean length of nap, all cases	82.3	76.8	5.5	1.2	4.6
Mean length of nap when nap was taken	85.4	82.9	2.5	1.0	2.5
Mean time for going to sleep	28.7	32.3	3.6	.8	4.5
Percentage naps missed	3.8	7.8	4.0	1.0	4.0

There appears to be, in the case of the nursery-school children studied, very little if any immediate relationship between the amount of morning outdoor play and either the length of nap or the length of time required to go to sleep for the afternoon nap. It is possible, however, that prolonged periods of staying indoors might bring different results. The children studied were never kept indoors for more than 5 consecutive days and only for 24 days during the entire experimental period. They did not attend school unless they were well and free from colds.

Influence of the Adult in Charge of the Nap Period. A situation existed which offered an opportunity to study the possible influence of the individual in charge of the children upon the afternoon nap. Two teachers divided the responsibility for the afternoon session at the nursery school, but only one of the two was present on any one afternoon. Each assumed charge for two or three afternoons weekly during an eight-months' period. The conditions otherwise remained the same, there was no change in the children's routine and the same assistant teacher was present for every afternoon throughout the week.[8]

[8]One assistant was present daily for the first four months of this study and another for the last four. Teacher X was in charge on Thursday and

Each child went to the sleeping room directly after finishing his luncheon. The assistant recorded the time when he lay down on his cot The teacher in charge did not enter the sleeping room until several children were already on their cots, usually about 15 minutes after the first child went to bed She left the sleeping room when the first children awakened and the assistant remained until all were awake. The time each subject went to sleep and awakened was recorded either by the teacher or by the assistant

One thousand nine hundred and sixty nap records were available, 962 for days X was in charge and 998 taken under the supervision of Y The average length of sleep, the average length of time elapsing between going to bed and going to sleep, and the percentage of naps missed were calculated for both sets of records and the differences compared Table 3 presents a summary of this comparison.

It will be noticed that the mean length of sleep was greater, the time required to go to sleep was less, and the percentage of naps missed was lower when the children were under the care of X than when Y was in charge All of these differences show statistical reliability Omitting the occasions when the children failed to go to sleep at all, the average length of nap is greater for X than for Y, although the difference here is less and within the limits of chance

No conclusive explanations can be offered for these differences Both teachers endeavored to use the same methods and frequently discussed all matters pertaining to the nap period Both joined the nursery-school staff at the same time, and were on the staff during the term previous to the period during which these records were made Their previous training for and experience in nursery-school teaching was similar During the term preceding the observational period Y had more experience than X in the conduct of the nap X, however, was daily in charge of the children during the morning session, while Y came at noon on specified days and remained for lunch and the entire afternoon This change from X to Y may have been a disturbing factor to some of the children, although they were all well acquainted with her. If the children who habitually missed naps happened to be absent more frequently on days X was in charge than on days Y was present, this might partially account for the above noted differences However, considering the records of the six cases who had missed at least 10 naps during the year, five missed a much lower percentage of naps under X than under Y Neither of the teachers nor the assistants had any knowledge of the future analysis of the records when they were made.

Reliability of the Method of Securing the Data. Closed eyes, cessation of

---

Friday during the first half of the observational period and on Monday, Tuesday, and Friday during the latter half Y was present on Monday, Tuesday, and Wednesday for the first four months and Wednesday and Thursday for the last four

movement, and regular breathing were the criteria for judging whether or not a child was asleep, and the reverse for waking up. In order to determine the extent of agreement between the different observers concerning these factors, the two teachers (X and Y) and the assistant (A) simultaneously kept the nap records for the group of children on one occasion. The resulting coefficients of reliability are as follows: For the time of going to sleep, +.99 between X and Y, +.99 between X and A, and +.98 between Y and A. For the time of awakening; +.99 between X and Y, +.98 between X and A, +.98 between Y and A.

It is evident that the afternoon sleep of young children may be very definitely influenced by the particular person in charge both in regard to the time required for the child to go to sleep and whether or not a nap is taken. Whatever the causes for this difference may be, they invite further investigation.

Afternoon Sleep and Rest before Luncheon. For a period of six weeks, 24 children were trained to take a 25-minute rest directly before luncheon, going to their cots in the sleeping rooms. Sleep was discouraged, although on rare occasions a child fell asleep and was awakened for luncheon. Following this six-weeks' period, for 19 weeks the group of children alternated daily between taking and omitting this rest, other factors remaining constant. Then followed a period of five weeks during which the daily rest was resumed.[4]

The averages for the length of nap during the above-described periods are as follows: days when no rest was taken, *70.4* min.; days when rest was included, *70.9* min.; control period (preliminary and final periods when daily rest was taken), *70.5* min. The averages for the length of time required to go to sleep are: non-rest days, *29.1* min.; rest days, *29.8* min.; control period, *30.8* min. These differences are very slight and unreliable, indicating that the omission or inclusion of a 25-minute rest period before luncheon has little or no effect on either the length of the afternoon sleep or the length of time elapsing between going to bed and falling asleep.

SUMMARY

1. Records were kept of the amount of time spent daily in outdoor play and of the length of the afternoon sleep of 30 nursery-school children.

2. There appears to be, in the case of the subjects studied, little if any relationship between the amount of time spent out of doors in the morning and either the length of the afternoon nap or the length of time required by the children to go to sleep.

3. The particular adult in charge of the children during the afternoon nap may influence both the length of time required by the children to go to sleep and the number of naps missed.

[4]These data were collected during an investigation by Mildred Larson on "The Effect of a Rest Period before Luncheon on the Noon Food Consumption of Nursery School Children."

4. The omission or inclusion of a 25-minute rest period before luncheon had no noticeable effect on the afternoon sleep of the children.

REFERENCES

1. FOSTER, J C, GOODENOUGH, F L, & ANDERSON, J. E The sleep of factors *J Genet. Psychol*, 1930, **38**, 114-126
2. SHERMAN, M Afternoon sleep of young children: some influencing factors. *J Genet. Psychol*, 1930, **38**, 114-126

University of Nebraska
Lincoln, Nebraska

A CRITICAL NOTE ON SOME RECENT RESEARCH ON MECHANICAL ABILITY

JOHN W COX

GENERAL SCOPE AND METHOD OF THE MINNESOTA AUTHORS

There has been published recently in America an account of a research on mechanical ability (2) It is the joint product of the psychological staff of the University of Minnesota, supported by many well-known American names, aided by a large body of graduate students, teachers, testers, and clerical assistants, and financed by the National Research Council. As befits its eminent origin, its production has been carried through on a most liberal scale

The general aims were, firstly, to construct an adequate test for measuring mechanical ability and, secondly, to determine the internal organization of mechanical ability and its relation to other psychological traits and to environmental conditions Mechanical ability was defined generally as "the ability to succeed in work of a mechanical nature" The scope of the research was, however, limited to a "working definition," viz, "that which enables a person to succeed in a definitely restricted range of vocational and trade school courses." In this, two aspects were recognized, viz., "the ability to succeed in the actual manipulation of tools and materials and the ability to secure information about tools, materials, and their uses"

In the light of this definition and a review of relevant literature, 26 tests were selected as providing a likely source from which a team of tests for adequately measuring mechanical ability might be developed These consisted of well-known "mechanical" tests, such as Stenquist's Mechanical Assembly and Picture Tests, Link's Spatial Relations Test, Cube Construction, and various formboard tests, and such "motor" tests as those of tapping, aiming, steadiness, balancing, and cancellation These, after suitable revision where necessary, together with intelligence tests, were tried out on 217 seventh- and eighth-grade boys attending a junior high school The validity of each test was determined by its correlation with the ratings assigned by shop teachers in the several shop courses which the boys attended (the chosen criteria of mechanical ability). Its reliability (i.e, cor-

relation with itself), where a retest was possible, was examined at great length, both in relation to the method of correlation employed and the method of scoring adopted

On the basis of these results seven of the tests (Stenquist Assembly and Picture, Link's Spatial Relations, Paper Form Board from Army Beta, Card Sorting, Packing Blocks, and Steadiness tests) were selected for revision with a view to increasing their reliability Revision consisted mainly in lengthening the tests Thus revised, they were given to a new group of 100 boys attending the same school—the experiment proper. In addition, the program was broadened to include such other indices of ability as academic success, previous mechanical experience, interests, agility and gymnasium success, measures of height, weight, and vital capacity, social and economic status and home influences; and greater attention was devoted to securing an adequate objective criterion This was narrowed down to 'mechanical ability as manipulation and information' and measured by (1) the quality of the work done in the various shops, (2) the quantity of the work done in relation to its quality, and (3) information about tools and materials and their uses

The correlations between the several tests (including the above-mentioned other indices of ability) and these criteria were taken as indicative of the extent to which the tests would predict mechanical ability Various test batteries for predicting mechanical ability (as measured by these criteria) were then constructed by the aid of Toops's multiple-ratio technique

The nature and organization of mechanical ability was next investigated by applying Spearman's methods of analyses to the intercorrelations of the tests and criteria To determine whether or not mechanical ability exists as a unique trait certain of the tests were compared in various ways (by the correlation method) with five other supposedly unique traits, viz., age, agility, intelligence, height, and weight

Their Results

The correlations between the 26 tests used in the preliminary experiments and the eight (shopwork) criteria selected as a measure of mechanical ability were generally low or negative (under 2). It was, however, possible to combine seven tests into a battery correlating 60 with the combined criteria of shop success and 07 with the intelligence test These formed the basis of the tests employed in the experiment proper As revised for this purpose, their reliability was raised to the neighborhood of 8- 9 The reliability of the various criteria adopted in the experiment proper also approximated to 8

In the experiment proper the correlations between the tests and the criteria were again generally low but better with 'quality' (19- 55) and 'information' (—04-.57) than with the 'quality-quantity' criterion (—01- 24). By combining five of the measures (Paper Form Board, academic grades,

Minnesota Assembly, Son's Mechanical Operations Questionnaire, and the Interest Analysis Blank) with the intelligence test a battery correlating .81 with the combined quality and information criteria was obtained Eight other batteries of varying composition were found to correlate 55-.73 with the quality criterion. To this exent the batteries were held to predict mechanical ability

The intercorrelations of the tests and other data collected were found to be low and suggested to these authors "that specific factors rather than a single general factor characterize mechanical ability" Their application (which we criticize later) of Spearman's methods of analysis leads them to conclude that the evidence "indicates rather definitely the presence of group factors. However, no claim is made that a general factor is disproved by this evidence."

The correlations of a selected test battery with the other 'supposedly unique traits' (mentioned above) were found to be low, while its correlation with the 'quality' criterion was relatively high (.61) This provides their main evidence for declaring mechanical ability to be a unique trait, although they add, somewhat confusingly, that intelligence has not been completely eliminated

Critical Examination of Results

Enough has been said to indicate the vast amount of painstaking labor that has gone into this research In addition to the work already mentioned, an inquiry was made into the effects of environment on, and the distribution of, mechanical ability (as measured by these tests) The work also contains complete instructions for using the tests, the complete scores made by all the subjects on all the tests and criteria, and an exhaustive bibliography Enough, too, has been said to indicate that useful results relating to the practical employment of the tests were secured.

When, however, we turn to the central problem of this research—the nature and organization of mechanical ability—we fear that the authors' plan of attack, and their claim to have discovered and demonstrated the 'uniqueness' of mechanical ability, are open to serious criticism

1 *The Nature of Mechanical Ability* Evidence that mechanical ability involved a special group-factor, together with the method of measuring this special ability, was published by the present writer some years ago (1). He has therefore carefully examined the evidence put forward in the present work His conclusion is that the book's main service to fundamental research in this field lies in the examples it provides of the type of error open to an unwary investigator

An unscientific attitude is adopted at the outset. Advance in science may begin psychologically either with a theory or with observed facts In the former case the theory must be tested by the subsequent observation of facts, in the latter a theory must be found to fit the observed facts. Logically,

both methods are, of course, complementary, the ultimate aim being both to discover the facts and to explain them by tested theories.

The research under review adopts neither of these methods. Nor does it follow a judicious combination of both. On the contrary, it begins by *assuming* the very point which it seeks to prove. We are told *at the outset* that "mechanical ability may be defined as the ability to do mechanical work," and that this "covers a wide range of activities from the invention of airplanes and the designing of bridges to the mere insertion of pieces of metal in a machine or the stamping of holes in leather." Such a proceeding might be excusable if it were intended merely to mark out the field to be surveyed. But this is clearly not the case, for the work is vitiated throughout by the preconception of this highly complex group of activities as a single trait. It appears in the widely differing nature of the tests and *criteria chosen to measure* the ability so defined, in the *forced interpretation* put upon the correlations to the neglect of other traits, and in the failure to observe the incongruity of finding no less than seven group factors in this supposedly unique trait. The overpowering influence of this preconception is seen when the results of a lengthy preliminary experiment, specially carried out to determine the best tests for 'mechanical ability,' are set aside in favor of the Stenquist Assembling Test "because of its previous reputation as the best all-around test of mechanical ability." It is, perhaps, this 'all-around' notion which contributes most to their undoing, for a *'unique' trait is not an 'all-around' affair.*

2. *The Measurement of the 'Unique Trait.'* Having oriented the whole research according to this preconceived idea of mechanical ability, the authors naturally attain to no clearer notion in the special test batteries which are said to measure it, for these batteries are built up according to this preconception and involve a similar hotch-potch of abilities—manual, mechanical, interests, academic grades, and even general intelligence. They leave us more confused as to the nature of the unique trait they are said to measure than the ill-defined notion with which the research sets out.

Such procedure leads to no objectively determined and demonstrated trait, but merely to the observation that from a group of tests arbitrarily labelled 'mechanical' there can be selected some which will involve to only a small degree the sort of ability (itself open to question) measured by the 'intelligence' test. It is clear that the error involved here is not merely one of inadequately describing or explaining the observed facts, but the more important one of failing to collect the necessary kind of facts.

3. *The Criterion of 'Uniqueness.'* Let us now turn to another error, and one concerned with the interpretation of the data collected rather than with *errors of procedure in collecting them. In a research which claims to have* discovered a unique trait, the question of how the tests which are said to measure this trait are related to other tests is fundamental. The demonstration of uniqueness must depend upon a comparison between test scores

Since these are numerical, and the comparison is made by correlation coefficients, the relations to be observed, and on which the demonstration must depend, are necessarily mathematical It is therefore essential to define clearly the mathematical condition which a 'uniquely determined' score must fulfil, and to formulate a definite and adequate criterion by which the fulfilment of this condition shall be judged Having been formulated, the criterion should, of course, be consistently applied

The criterion proposed by the authors appears to conform to none of these conditions Their criterion for uniqueness is this "positive correlation with at least one criterion of human endeavor" combined with "correlations as low as possible with all other unique traits" 'Positive correlations' and 'correlations as low as possible' are very inexact expressions As thus stated, they provide no rigidly defined standard such as is required in a criterion Moreover, even if we proceed to define these more clearly, the criterion remains inadequate, for positive correlation is merely indicative of something in common between the correlated items For uniqueness, the "positive" correlation must attain theoretically to unity, and the "low as possible," to zero, moreover, the "criterion of human endeavor" would itself need to be unique Such standards are never approached in the research under review The very illusive nature of the factors which determine these criteria (ability in the workshop, interests, knowledge of tools and materials, information, and the like) is a good reason for seeking to discover unique traits

4. *The Application of the Criterion* Further, the criterion adopted was not consistently employed On the results of the preliminary experiments for devising a suitable battery, certain tests were selected not, as might be expected, on account of their high correlation with the criteria of mechanical ability, but for their low correlation with 'intelligence.' Finding (one supposes) that these selected tests, now alleged to be unique with respect to intelligence, are not in every case sufficiently good tests of 'mechanical' ability, the authors proceed to select a new battery based on correlations with the criteria of mechanical ability. In so doing, the requirement of low correlation with 'intelligence' recedes into the background, new tests are introduced, and the batteries finally selected for predicting mechanical ability as a unique trait are heavily weighted with 'intelligence' This peculiar result seems traceable to failure to distinguish between the causes, or factors, which determine the test scores (the traits one seeks to isolate) and the scores themselves

5 *Misunderstanding of the Two-Factor Theory.* This brings us to the most serious error—the misunderstanding of the criterion which does make this important distinction and which was long ago formulated by Professor Spearman Our first inkling of this confusion is an historical error It is stated that the theory of two factors was not definitely advanced until 1914, whereas in fact it was formulated ten years earlier (3, p 284).

Hard upon this follows a misstatement of the theory itself "The general principle," we are told, "is simple enough If the correlations between the tests are zero, or very low, the capacities tested are independent and specific, if the correlations are very high a general factor has been at work" This bears a dangerous similarity to the authors' own 'criterion,' but with 'capacities' and 'factors' substituted for 'unique traits' Was ever a 'simple' principle so entirely misunderstood? The merest acquaintance with Spearman's theory should suffice to show that it is not based on the absolute magnitudes of the correlations but on their ratios to one another

6 *Attempted Application of the Two-Factor Theory* When the authors attempt to test their own data by this wrongly conceived criterion they sink yet deeper into the mire They object that the inter-columnar correlations do not average the 1 00 required by the theory But they have used a wrong formula (4, p 82, formula 8), and under conditions where even the right one would not be applicable (4, p 56) Moreover, this inter-columnar form of the criterion has long been obsolete, having been replaced by the more convenient 'tetrad-difference' form (5, pp 138-140)

With the 'inter-columnar' criterion in mind, the authors find that eight of their measures could be arranged to form seven hierarchies of four tests each. These are then shown to be perfect by the tetrad-difference criterion Hence they conclude that their mechanical ability batteries involve seven group factors Nowhere is it shown how far these correlational differences may result from differences in the degree of 'intelligence' involved in the various measures, nor how far to the intrusion of non-cognitive influences arising out of environment—especially in such measures as academic grades, information, quality of work done in the shops, and mechanical operations (recorded as having been done about the house) All of these play a larger part in the Minnesota measures

7 *Relation to the Reviewer's Own Finding* An examination of the data in relation to the reviewer's own work on mechanical aptitude must be reserved for a later paper It may be said, however, that the presence of a 'mechanical' factor in certain portions of the Minnesota data would accord with his own findings (1), long since published. He, however, was obliged to relinquish the assembling type of test (on which the Minnesota authors largely rely) and devise special cognitive tests Moreover, in a subsequent research soon to be published he finds that the manipulative aspect of assembling work involves an ability not to be confused with the 'mechanical' factor which he found to underlie the inventive aspect stressed by his tests. Such distinctions do not enter into the Minnesota batteries, which probably include these and many other factors. Neither could the doctrine of unique traits as described by these authors lead to fundamental distinctions of this kind

To discuss this doctrine would carry us beyond the scope of the present article. It must suffice to say that, from the vague account given, their

'unique traits' appear to be not psychological entities, but mathematical quantities (the measures or test scores) from which something else (human endeavors or social reactions) may be predicted (2, p 245, footnote) If we attempt to interpret them psychologically, it is difficult to see how they differ from the 'factors' of Professor Spearman and his pupils, except that the important distinction between the whole test score and that part of it determined by the factor is entirely disregarded Nowhere in the research is any proof given of the existence of a 'unique trait' in the only intelligible sense of this term, namely, a quality of mind which functions as a unitary whole and in complete independence of other mental qualities. And nowhere is there any attempt at a psychological analysis of such a trait.

REFERENCES

1 Cox, J W Mechanical aptitude its existence, nature and measurement London Methuen, 1928 Pp iii+209 (First reported in a paper read before the Education and Psychology Sections of the British Association in 1925.)

2. PATERSON, D. G, ELLIOTT, R M, ANDERSON, L D, TOOPS, H A, & HEIDBREDER, E Minnesota mechanical aptitude tests. Minneapolis: Univ Minn Press, 1930 Pp xxii+586

3 SPEARMAN, C 'General intelligence' objectively determined and measured *Amer J Psychol*, 1904, 15, 201-293

4. ——— General ability, its existence and nature *Brit J Psychol*, 1912, 5, 51-84

5. ——— The abilities of man London Macmillan, 1927. Pp. xxiii+415.

University of London
London, England

APPARATUS

A LIFE-SIZE ALLEY MAZE FOR CHILDREN

F Davis and M Batalla[1]

This note is for the purpose of describing a life-size alley-type maze which was erected in 1930 on the grounds of the University of California Institute of Child Welfare

Construction of the Maze Within a fenced enclosure of approximately 400 square feet, a wooden platform was constructed, measuring 16′ by 16′ This was pivoted at the center upon a length of 2½″ iron pipe, which was screwed into a steel plate and set in concrete The pivot construction was intended to permit the maze, when set up on the platform, to be rotated and placed in a new orientation, without dismantling it, the purpose of rotation is to control orientational cues which may be present in the environment and which may influence the learning of certain types of mazes Removable wooden supports, at each of the four corners, maintain the platform in a level position, approximately one foot above the ground

The walls of the maze were made of tongue and groove redwood, and were constructed in sections, each section being 3′ wide and 4½′ high The sections were joined by fitting them into grooves in the sides of pine posts, each 4″x4″x4½′. When the walls were completely assembled structural rigidity of the whole maze was obtained by fastening lengths of pine to the tops of the posts, these pine strips, running the full length and width of the maze, were held in place by iron rods. Thus the construction permits of speedy assembling and disassembling of the maze units whenever it is desired to change the pattern for different experiments, or to clear the platform for other uses

The doors were made of tongue and groove white pine, measuring 2½x4½′, and were hinged at top and bottom with spring hinges (see Figure 2) When the doors are pushed open, the springs return them to the closed position They are prevented from swinging beyond this closed position by stops extending the full length of the post In addition to acting as blocks when the doors swung back, these stops also serve to hold the closed doors in the proper position to permit the solenoid lock mechanisms to work reliably When unlocked, the doors could be opened from the outside by pulling on handles attached to them for that purpose

Door-Locking Mechanisms Locks for the doors were made as follows

[1]Acknowledgments are due to Professor Harold E Jones for guidance in the planning and construction of the maze

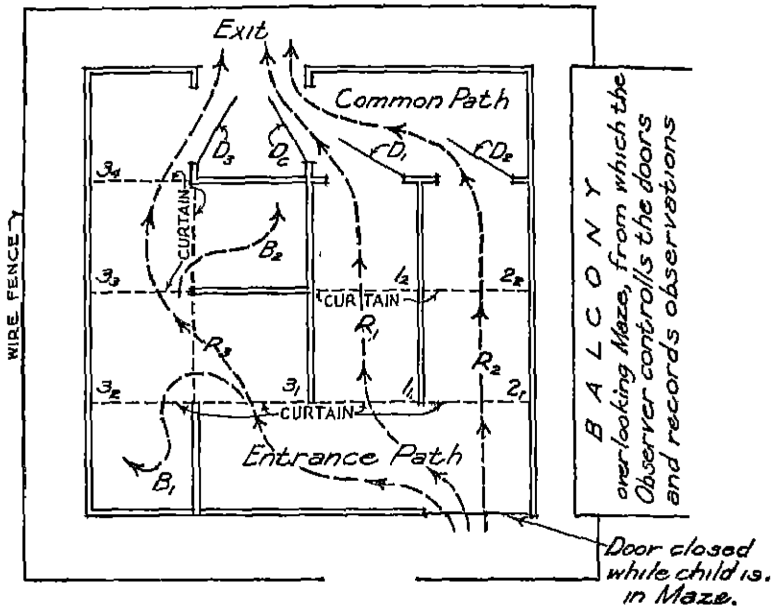


FIGURE 1
PLAN VIEW OF MAZE

1 Solenoids were constructed and were sunk into the posts opposite the closed doors at approximately 60-degree angles

2. In line with the solenoids brass cylinders ¾" in diameter and 1" long were sunk in holes bored in the door edges opposite the posts

3 A soft iron bolt, ¼" in diameter and 3½" long, is drawn up into the solenoid when the circuit is closed for a given door, thus unlocking the door. When the same connection is broken, the bolt slides back through gravity into a brass cylinder in the door, locking the door once more until the current is again turned on at the experimenter's switch (See Figures 2 and 3.)

Figure 4 shows how the solenoid locks are connected with the switchboard controls, the latter being located on a balcony overlooking the maze It also shows the buzzer connections, by means of which the experimenter is able to communicate with an assistant who starts the child in the maze, in accordance with codified signals.

Maze Plan The maze pattern used in our first series of experiments is illustrated in Figure 1. It is similar in its general plan to the "final common path" maze suggested by Tolman for use with rats, and first used by

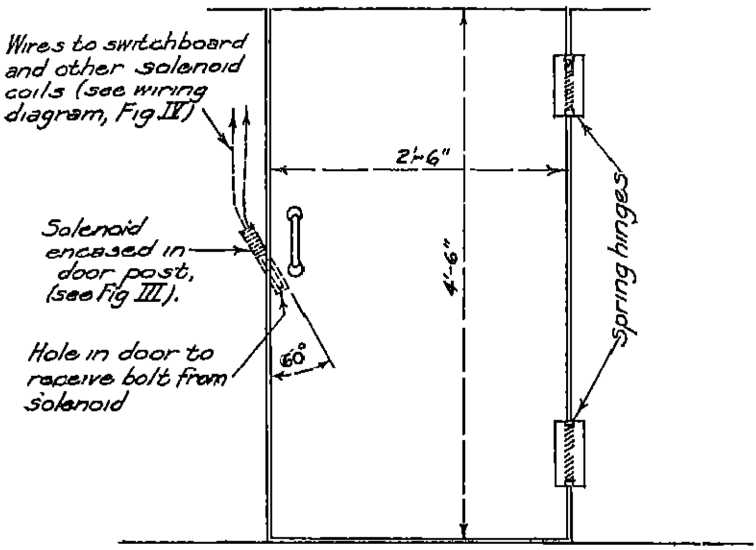


FIGURE 2
SKETCH OF DOOR, SHOWING SOLENOID COIL IN POSITION

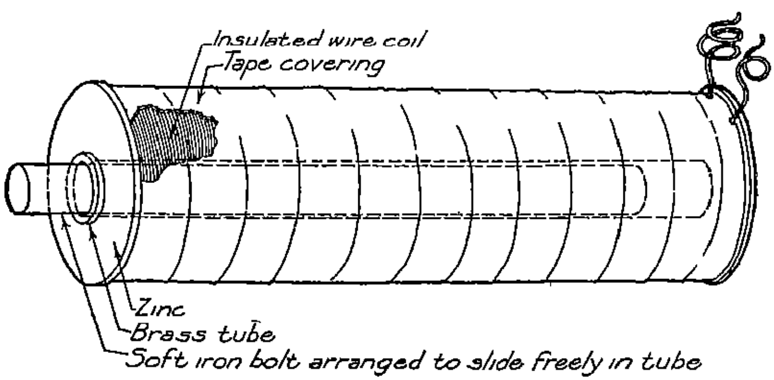


FIGURE 3
DETAIL OF SOLENOID
(Reduced to 4/8 7/16)

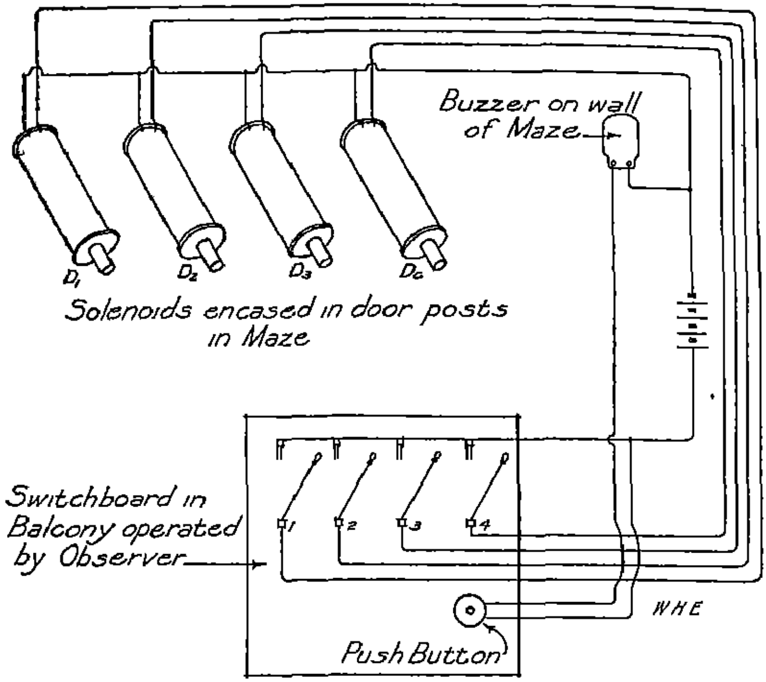


FIGURE 4
WIRING DIAGRAM FOR ELECTRIC CONTROL OF DOORS IN MAZE

Hsiao (1) in his study, "An Experimental Study of Rats' Behavior Within a Spatial Complex."

The maze illustrated presents three possible routes from entrance to exit The first route, *R1*, leads to the exit via doors marked *D1* and *Dc* The second route, *R2*, goes from the entrance door straight ahead down the right-hand alley to *D2*, and then out through *Dc* It will be noted that *R1* and *R2* both lead into a final common path, with a common door, *Dc*, opening out from this common alley and leading to the exit The third route, *R3*, is longer than either *R1* or *R2*, and leads to the exit directly through a door of its own, *D3*, not through *Dc* It differs further from the other two routes in that it has two blind alleys, *B1* and *B2* (The curtains indicated in the diagram were strips of blue denim hung from post to post wherever there was neither door nor partition Their purpose was to limit the child's visual inspection of the maze)

Records From his position behind a screen in an elevated balcony, the

experimenter is able (unseen by the child) not merely to control the maze pathways but also to record in code a detailed account of the alleys entered, of retracing, and of individual differences in method and attitude, as shown in exploratory tendencies, active trial and error, perseveration, hesitant or tentative behavior, etc Time is also taken by stop-watch, and on special occasions moving picture records have been obtained, presenting a bird's-eye view of characteristic behavior patterns in the maze The assistant who brings the child to the maze gives the necessary verbal instructions and starts him according to signal This assistant is unable to see the child after the maze has been entered, but vocalizations can be heard and recorded; an account is also taken of relevant features of the child's behavior before and after running the maze

Merits and Defects of the Life-Size Maze Two minor difficulties were encountered in the operation of the maze itself The technique of switchboard operation, in unlocking and locking the doors, requires a degree of alertness in the operator first, to see that the right door is unlocked before the child attempts to open it, and, secondly, to make sure that the connection is not "broken" too soon after the child has passed through the door, since if the circuit has been reopened before the door has closed firmly in position the bolt will fall to the floor of the maze and will have to be replaced manually before the experiment can continue Very little practice, however, suffices to assure the degree of skill necessary to prevent such mishaps An improvement in the wiring would consist in the introduction of an additional circuit for each door, which would maintain an automatic clutch of the solenoid as long as the door is open. The other difficulty referred to has to do with the current supply for the solenoids We used a six-volt storage battery, which was satisfactory as long as it was kept fully charged, the drain upon it was sufficiently great to necessitate recharging about once every four days Line current, with a step-down transformer, would be a more satisfactory source of supply if it were readily available

Because of the demountable unit construction of the maze, a multiple-T or other type of maze can in a relatively short period be substituted for the pattern above described Perhaps the chief merit of a life-size maze, for the study of children's learning, *is that to a certain degree at least* such a situation is self-motivating The maze is an object of interest to all the children, and they vie with one another for the opportunity to enter it. A year's experience indicates that this equipment provides a flexible and effective method of studying the maze behavior of children between the ages of two and eight years

Reference

1 Hsiao, H. H An experimental study of rats' behavior within a spatial complex *Univ Calif Publ Psychol*, 1929, **4**, No 4, 57-70

Institute of Child Welfare
University of California
Berkeley, California

BOOKS

EDITED BY

EDMUND S CONKLIN
SANTE DE SANCTIS
A. R LURIA
GODFREY THOMSON
C J WARDEN

THE CHILD IN THE THIRD DECADE

A Review

The Editor of the International University Series in Psychology has conceived and called into being some amazingly useful compendia, of which the present volume[1] is the most recent. Like most of its predecessors in the series—particularly *Psychologies of 1925, The Foundations of Experimental Psychology,* and *Psychologies of 1930*—it is sure to find almost at once a place on the active shelf of every psychologist whose interests and understanding extend beyond the borders of his research specialty Not alone as a documented "handbook" for ready consultation, but as a book to be read ponderingly, to define frontiers of knowledge, and to point to unexplored areas, the volume is a major contribution It is to be hoped that within five years at the most another similar handbook may take its place beside this one, that, in fact, a series may be projected which will record and consolidate at regular intervals the gains made in the rapidly moving field of child psychology

Twenty-two authors have undertaken the task of compiling a volume "treating the entire field of child psychology which could be useful to students already acquainted with psychology and already expert in child psychology" The work has been so divided that each author summarized a zone in which one of his major research interests has lain—thus, for example, Gesell on "The Developmental Psychology of Twins," Peterson on "Learning," C. Buhler on "Social Behavior," Goodenough on "Children's Drawings," Terman on "Gifted Children," etc

In general, the contributions are written in a heavy, condensed handbook style, which is perhaps unavoidable when a great store of information is to be assembled without exorbitant use of space Practice differs

[1] Murchison, C. [Ed] A Handbook of Child Psychology Worcester. Clark Univ Press, London. Oxford Univ Press, 1931 Pp xii+711 $5 00

greatly among the authors with respect to their presentation and interpretation of data Some give conclusions supported only by names and reference numbers, others give comprehensive summaries of data, but very little interpretation, still others give facts and interpret them also, and one or two concern themselves only with the applied field Franzen used to say that there ought to be a law against presenting anything in published research reports except uninterpreted statistical data Some of the *Handbook* chapters appear to have been written from such a slant and some have not been, the former give the impression of greater scientific accuracy, but the latter are the most interesting.

The authors differ, also, in the thoroughness with which they have combed their fields A few frankly present condensed accounts of their original work rather than summaries of studies in which their own work may constitute a link. Examples are Isaacs, whose chapter tells of a three-year experiment with "an environment optimal for mental growth," and Piaget, whose original approach to the child's conceptions of realism, animism, and artificialism provides material for one of the most interesting chapters in the entire book Many authors, on the other hand, give summaries which the reviewer believes to be exhaustive for at least the most significant experimental contributions From 80 to 150 citations are not unusual in such chapters A few "summary" chapters, however, show evidence of haste in compilation, and leave impressions with the reader which are incomplete, because important studies, particularly of the past two or three years, are left out, or because significant aspects of the special fields are not considered at all

Types of Contribution

In addition to the few contributions which give in abbreviated form the results of the authors' own researches, the chapters seem to group themselves in several other clearly defined classes First, there are the ones which present primarily *points of view,* discussions of problems, methods of work, and the proper provinces of child psychology, or "best" judgments as to correct child training We may next distinguish the chapters built about a definite field of child behavior or child activity—social development, learning, dream life, language, play life, etc To this division the largest single group of contributions belong Again, we may consider together the chapters which are concerned with certain important categories of childhood—e g , the gifted child, the feebleminded child, the eidetic child, the primitive child Finally, we may group the few chapters which are concerned chiefly with the forces influencing child development

While the divisions necessarily overlap somewhat in their content, they are sharply differentiated in their organization of material It is a stimulating experience for the reader to do his thinking now in terms of the

special modes of child behavior, now in terms of special kinds of children with all the complexities which make up their total personalities, now in terms of the natural and artificial influences which modify the human organism for better or for worse

Chapters Presenting Points of View

John E Anderson opens the volume with a chapter entitled "The Methods of Child Psychology." Analyzing the field of child psychology in terms of the scientific method, he traces the development of view points and techniques from the time of Plato through the vigorous speculations of Rousseau and Locke, the halcyon days of baby diaries, the deluge of questionnaires of which Stanley Hall was the inspiration, into the present era of measurement and experiment. Discussing contemporary trends in child psychology, the author distinguishes six useful sources of material, and twelve methods of approach An interesting discrimination is made by him between those dubious questionnaires asking for expressions of opinion, and questionnaires asking for records of fact The section on "control by statistical devices" (the last of the twelve methods of approach) is not altogether sound, particularly in its consideration of the partial correlation technique The chapter ends with a very quotable generalization "Scientific sin consists not so much in the use of a particular method as in the failure to use a more adequate method for the problem in hand when such a method is available"

"Psychoanalysis of the Child," by Anna Freud, makes the modest claim that "Psychoanalysis does not permit itself to be ranged with other conceptions: it refuses to be put on an equal basis with them Instead of taking its place beside the others, it usurps the function of writing the entire textbook of child psychology on the basis of its own discoveries" Excusing herself from giving research results because of lack of space, the author sketches the child's sexual (emotional) development from infancy through puberty as viewed by the psychoanalytic school The chapter, highly condensed and lacking illustrative material, will probably be all but unintelligible even to well-trained psychologists, with the exception of those who are already steeped in the jargon of psychoanalytic literature

Helen T. Woolley's chapter, "Eating, Sleeping, and Elimination," is included by the reviewer under "points of view," even though the title suggests grouping it with "fields of child behavior," because the point of view of the parental education specialist is the dominant note While it is true that several research studies upon eating, sleeping, and elimination are mentioned, this is incidental to the manifest purpose of the chapter—to provide a manual of child training in three important sets of habits Since there is probably no aspect of child psychology so thoroughly explored as to provide a completely validated groundwork for pedagogy, it is usually

necessary in formulating a plan of training to resort to enlightened judgment and to practical common sense in the areas where research indicates no clear line of action. This Professor Woolley does admirably When even the "authorities" fail her, she is nearly always able to bring up some pat and convincing illustration from her own experience as a mother. But some authorities might question her dictum that "the child should go to sleep in the same room, in the same bed, dressed in the same clothes, and put to bed by the same person every time he is started upon a night's sleep or nap." What price motherhood!—what risk, indeed, of establishing some of the infantile fixations which the psychoanalysts and behaviorists alike decry!

Chapters Presenting Original Research

There are only two chapters which fall distinctly under this heading and which, therefore, present but a segment of a field instead of a broad survey Many readers may agree with the reviewer in voting these two chapters the most interesting in the entire book—"The Experimental Construction of an Environment Optimal for Mental Growth," by Susan Isaacs, and "Children's Philosophies," by Piaget. The first of these gives a fascinating and detailed account of an experimental "progressive" school conducted for three years by the author in England The second describes some of the author's experiments upon children's conceptual thinking, particularly the development of the child's conceptions of realism, animism, and artificialism Valuable though the two contributions are, it is questionable whether the appropriate place for them is in a reference handbook. More might have been gained by devoting the space to one or two of the fields in which the handbook is admittedly incomplete

Chapters Presenting Fields of Child Behavior

It is in this group of chapters that most of the surveys are almost encyclopedic in organization and in thoroughness

Beth L. Wellman ("Physical Growth and Motor Development and their Relation to Mental Development in Children") presents a résumé of experimental work upon growth by age and sex of physical and motor traits, then permutes mental, physical, and motor development into the three possible combinations of each with each and summarizes the data bearing upon their interrelationships. Little interpretation of the significance of the material is given—much of the chapter, in fact, would not suffer if thrown into a series of tables with brief explanatory legends

Dorothea McCarthy ("Language Development") is concerned chiefly with the development of language in the preschool child, since the early ages represent the period of most rapid development, and the normal child past five is already using all parts of speech, all inflections, and all sentence

forms Beginning with prelinguistic utterances, the author traces the developmental stages of speech, and then gives well-documented accounts of the growth of vocabulary, the development of the sentence, the functions of language, and the relation of language development to other factors such as socio-economic status, twinship, intelligence, and age of associates Speech pathology is not considered in this chapter A valuable feature is a brief section outlining the directions which further research may most profitably take

In a chapter on "Learning in Children," Peterson's method has been to single out a few representative research studies to summarize in elaborate detail, referring to supporting or conflicting studies only by author's name or very briefly with respect to results. A wide variety of learning problems are discussed, infant coordinations, maturity versus training, acquisition of skill in young children, effects of learning conditions, transference, and various types of learning—associative learning, rational learning, etc. Strange to say, almost nothing is presented upon the possibilities and limitations of training in relation to mental endowment, a field in which much crucial experimental work has been done during the past five or six years

Charlotte Buhler's chapter on "The Social Behavior of the Child" is exceptional for its orderly arrangement, thoroughness, and clarity It is also one of the few chapters to cover its field systematically both from the longitudinal view of successive developmental stages and the cross-section view of individual differences Normative studies from infancy to puberty are presented, and so skillfully integrated that a well-rounded picture emerges of the normal child's social development—his growing interest in group play, his responses to various types of social situations, his early and later periods of negativism, his readjustment of social values at adolescence. Various social types of children are also well described Some question remains in the mind of the reviewer as to the soundness of the author's distinction between problem children whose behavior is *exaggerated* and those whose behavior *deviates*. The chapter ends with an excellent summary of experimental data upon the relation of social behavior to life situations such as family constellations, institutional care, step and foster relationships, and the social milieu

Vernon Jones, in his chapter on "Children's Morals," contrasts moral development with intellectual development, the first reflecting "right or wrong motives or right or wrong senses of values," and the second involving "responses which, on the basis of existing knowledge, may be pronounced to be successes or failures" He proceeds to a discussion of individual differences represented in continuous distributions of such traits as honesty, cooperation, generosity, persistence, self-control, etc, describes the test, rating, clinical and social criteria by which such traits are measured, and presents six pages of tabulated data regarding character tests now

available Data on the influence of various factors upon moral development are then summarized—nature factors including intelligence, age, sex, race, and "native tendencies to action," and nurture factors including home, associates, Sunday School, day school, and recreational activities These sections provide an exceedingly fine digest of recent research results, yet there are certain regretable omissions In the section on age, for example, test results showing little effect of age are cited, but the striking facts regarding age trends in the incidence of juvenile delinquency are not mentioned In a section on inheritance no mention is made of Lange's arresting study of criminality in fraternal and identical twins, although that study, with the one of May and Hartshorne on deception in siblings reared in orphanages, gives us almost the only crucial evidence we have upon the heredity of moral tendencies. The May and Hartshorne sibling study is cited, but the essential fact that these sibling pairs were reared in orphanages is not mentioned The chapter ends with an excellent discussion of learning experiments in the field of moral behavior.

The chapter on "Children's Drawings" by Florence L. Goodenough represents a remarkably fine integration of experimental work and interpretation. With facile familiarity the author gathers up the material of past decades, linking it to the most recent contributions She also draws more liberally upon foreign data than do many of the American contributors to the *Handbook*. That the drawings of young children are a universal language for the expression of ideas rather than for aesthetic expression is the author's well-supported contention The growth of perception in children is illuminated by the fact that the developmental stages in representative drawing are the same for children of all races

"Children's Plays, Games, and Amusements," by Helen Marshall, opens with a brief discussion of play as a biological tendency Classic compilations of plays and games are cited which reveal the play-life of childhood to "be strikingly independent of temporal or racial barriers." Sections follow which consider the influence of age, sex, intelligence, environment, and race upon play. A page devoted to the topic "Playgrounds and Organized Recreation" lists the organized facilities provided by cities, and comments on the paucity of opportunities in rural communities The chapter as a whole deals almost exclusively with children's games, and little, if at all, with the wider aspects of children's plays and amusements Leisure time activities such as reading, constructing, collecting, etc, are not considered

Kimmins' contribution on "Children's Dreams" is in the main a summary of his own data collected in the London schools, and would have been classed by the reviewer in the group of "original" chapters were it not for the informing brief supplementary reviews of theories of the nature of the dream and day-dream, the child's conception of the dream, and dream control Taking the view that the dreams of normal children "let in a flood

of light as to the temperament and mental make-up of the child," the author proceeds to classify dream types according to their normal appearance among children of the two sexes at various ages. Interesting sections follow upon the dreams of children in industrial schools, and of children who are deaf or blind.

"Special Gifts and Special Deficiencies" which are relatively uncorrelated with "*g*," or general intelligence, are discussed by Leta S Hollingworth Evidence is presented for the independence of musical, drawing and mechanical abilities Achievement in such school subjects as reading, spelling, or arithmetic, while usually found to be closely in line with "*g*," occasionally lags so far behind "*g*" that special remedial methods become necessary Ability to lead and govern people is conceived as a combination of physical and temperamental traits with intelligence in an "optimum section of the curve for *g*"

This chapter, one of the briefest in the entire volume, does little more than hit the high spots of a fascinating field of child psychology. Precise data regarding the inter-relationships of general and special abilities are not summarized at all, the developmental aspects of special abilities are scarcely touched upon, and recent experimental work leading toward theoretical interpretations (such as the relationship of lateral dominance to reading deficiency) is ignored The question may be raised as to whether the Editor's suggestions to contributors were sufficiently definite and clear, for the chapter gives the impression that the author felt an unnecessary pressure to select and condense

The Psychology of Children of Certain Typical Groups

If all the fields of child psychology had been covered by *Handbook* contributions, and if all the authors had provided exhaustive surveys of these fields, it is fair to say that the chapters on child types would represent chiefly a reclassification of data already summarized As the chapters actually appear, however, there is comparatively little overlapping between those grouped under the present heading and those previously reviewed The summaries of the special fields deal for the most part with the normative, developmental aspects of the field Among the sections which consider individual variations also, only a very few present with any degree of completeness the relation of such variations to the child's status in other important respects, such as intelligence The five chapters here classified, therefore, which include discussions of the gifted child and the feebleminded child, contribute material which is largely unique

Terman's chapter on "The Gifted Child" discusses those children who fall in approximately the top half of one per cent in intelligence Facts are summarized regarding their physical, intellectual, and personality traits, social origin, scholastic achievement, educational needs, and longitudinal

development over a period of years. The author emphasizes the fact that wide deviations from the central tendencies in all aspects of development are to be found among individual cases of gifted children.

Despite the fact that his material is more highly condensed than that of the majority of the *Handbook* chapters, Terman's presentation is extremely readable. This is partly because he largely dispenses with citations and summaries of specific data, weaving instead a skillfully integrated narrative discourse. The reader probably gets more from a first reading of such a chapter than from a summary of the more usual kind, although the method limits the usefulness of the contribution as reference material.

In a chapter on "Feeblemindedness," Pintner gives a comprehensive, well-documented review which includes the history of scientific study of the feebleminded, the incidence of feeblemindedness in the general population, clinical types, mental and physical characteristics, causes, and the care and control of the feebleminded. A valuable feature of this chapter is a one-page summary which concludes it. An omission which the reader is likely to regret is a discussion of the personality traits of the feebleminded, and particularly of the relationship of feeblemindedness to delinquency.

Kluver contributes a chapter on "The Eidetic Child." His presentation is extremely technical. It consists in large part of critical comments upon methods and conclusions to which he refers in terms so sketchy as to connote little meaning for the reader not thoroughly conversant with the original sources.

Margaret Mead, who never fails to provoke thought, approaches her chapter on "The Primitive Child" in a manner unique from that of any other *Handbook* contribution. The major portion of the chapter consists of a discussion of research problems important to the entire field of child psychology, whose solutions appear possible through data obtainable from a study of children in primitive societies. Such data as have already been secured (much by the author of the chapter) are grouped according to the problems which they suggest and illuminate—a form of organization which demands a particularly high order of selective and relational thinking. Studies of the effect of weaning, sleep habits, and demonstrative affection; studies of the fixation of emotional responses, studies of the development of language and thought, of musical and artistic capacities—these are only a few of the excellent problems proposed by Dr. Mead. One difficulty, however, which the author largely ignores in this chapter, as well as in her past publications of original research, is that of probable native differences between the members of different primitive groups. Not always is it safe to attribute differences in child development to differences in social culture.

"The Developmental Psychology of Twins" by Gesell opens with a biological orientation, and then proceeds to sections on behavior correspondence, handedness, psychopathological correspondence, including Mongolism

and insanity, studies using "co-twin control," psychometric studies of twin groups, and the nature-nurture problem in relation to twin correspondence. The arrangement and sequence of material is unsystematic, and critical evaluation of the statistical studies of twins is not attempted The chapter, nevertheless, contains enough well-summarized information to constitute a very useful contribution

FORCES INFLUENCING CHILD DEVELOPMENT

The chapters falling clearly under this heading are only three in number. Many will wish that more space had been devoted to such material, for it bears the same relation to descriptive child psychology that astrophysics bears to a celestial atlas While many of the *Handbook* chapters reviewed in other divisions contain sections which discuss the relation of a certain aspect of child development to various determining factors, few are as thorough on this point as the importance of the subject deserves

Mary C Jones writes a chapter on "The Conditioning of Children's Emotions" With editorial comments interspersed, she summarizes the available studies on conditioning, re-conditioning, and unconditioning, and discusses the problems concerned with the generalization of responses, intensity of stimuli, and substitution potentiality The point of view of the author is cautious, and does not embrace the startling inferences often drawn from such data as she reviews by those who go in for Behaviorism as a cult

Lewin's chapter on "Environmental Forces in Child Behavior and Development" is disappointing for what it leaves out, though stimulating for what it puts in The discourse is chiefly theoretical, though illustrated by citations from experimental research The entire chapter is really an elaboration of the fact that "dynamic" environment has a differential effect depending on the momentary condition of the individual Little attention is given to the many recent statistical studies of the effect of nurture upon intelligence and achievement, the author taking the questionable view that "the concepts of the average child and of the average situation are abstractions that have no utility whatever for the investigation of dynamics"

In a chapter on "Order of Birth in Relation to the Development of the Child," H. E Jones summarizes the available information regarding birth order in relation to intelligence, language, emotions, and delinquency The evidence is conflicting, and in general the effects are rather slight in so far as statistical studies are able to show An important contribution of the chapter consists of a thorough, critical discussion of the snares which beset the worker in this field. Many of the discrepancies in published results can be accounted for by various disturbing factors such as incompleteness of families, effect of mother's age, varying sex ratio, differential mortality, and changing birth-rate

In conclusion, the reviewer would characterize the book as an immensely

valuable source of up-to-date information The chapters are uneven in quality, although the prevailing level of excellence is high. Few, if any, of the chapters are so comprehensive that a person contemplating a research problem would be safe in accepting their bibliographies without further exploration of the fields represented, though most of the chapters do succeed in giving a picture that is full and informing

In the Preface the Editor admits the difficulty of determining the limits to be covered in the *Handbook*, and mentions particularly the omission of a treatment of the delinquent child In the judgment of the reviewer this omission is to be regretted, since the control of delinquency is one of the major problems confronted by child psychologists in the clinical field Among other fields which are treated inadequately, and which future *Handbooks* may find space to include, are

Development of sensation and perception
Sensory defects (the blind child, the deaf child)
Native tendencies to action
Heredity and environment as determiners of individual differences
Nervous and mental disorders (problem children)
Interests
The rate and duration of mental growth

BARBARA S BURKS

Institute of Child Welfare
University of California
Berkeley, California

BOOKS RECEIVED

BEATLEY, B Achievement in the junior high school (*Harvard Stud Educ*, Vol 18) Cambridge, Mass Harvard Univ Press, 1932. Pp xvi+92 $2 00

BOOK, W F Economy and technique of learning Boston Heath, 1932. Pp x+534 $2 00

DOCKERAY, F C General psychology New York Prentice-Hall, 1932. Pp xxii+581. $3 50

DOLL, E A, PHELPS, W M, & MELCHER, R T Mental deficiency due to birth injuries New York. Macmillan, 1932 Pp xvi+289 $4 50.

LUECK, W R The arithmetical and algebraic disabilities of students pursuing first year college physics *Univ Iowa Stud Stud Educ*, 1932, 8, No 1 (New Ser, No 236) Pp 48

LUH, C W Word association in Chinese children *Yenching Stud Psychol.*, 1932, No 1. Pp. 66 (English abs, pp 7)

MILLER, D. W An orientation in educational psychology Boston Badger 1932 Pp 233 $2 00.

VALENTINE, C W The reliability of examinations an equiry. London, Univ London Press, 1932 Pp 196

VALENTINE, W L A psychology laboratory manual New York: Prentice-Hall, 1932 Pp vi+285 $2 00

$7 00 per volume
Single numbers $4 00

QUARTERLY

December, 1932
Volume XLI, Number 2

Two volumes per year

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

DECEMBER, 1932

Chaque article est suivi d'un résumé en français
Jedem Artikel wird ein Referat auf deutsch folgen

THE EFFECT OF INACTIVITY PRODUCED BY COLD UPON LEARNING AND RETENTION IN THE COCKROACH, *BLATELLA GERMANICA* . 253
By Walter S Hunter

THE PERSONAL EQUATION IN RATINGS I AN EXPERIMENTAL DETERMINATION . 267
By Herbert S Conrad

A STUDY OF CERTAIN SELECTIVE FACTORS INFLUENCING PREDICTION OF THE MENTAL STATUS OF ADOPTED CHILDREN, OR ADOPTED CHILDREN IN NATURE-NURTURE RESEARCH 294
By Alice M Leahy

A STUDY IN RATING TECHNIQUE WITH SPECIAL REFERENCE TO ACTIVITY IN PRESCHOOL CHILDREN 330
By Helen Lois Koch and Helen Sirfit

A FURTHER STUDY OF THE INITIAL MAZE BEHAVIOR OF RATS . 358
By Wayne Dennis, R H Henneman, and W B Varner

THE INVENTIVE GENIUS OF THE CHILD . 369
By Rosa Katz

(OVER)

Worcester, Massachusetts

Entered as second-class matter August 3, 1897, at the post-office at Worcester, Mass, under Act of March 3, 1879

THE IMAGINARY PLAYMATE A QUESTIONNAIRE STUDY 380
By E B Hurlock and M Burstein

SOME EXPERIMENTS WITH KITTENS ON THE SIMPLE ALTERNATION PROBLEM . . 393
By Audrey M Shuey

LEARNING VERSUS SKILL IN RATS 406
By Joseph G Yoshioka

AGE AND SEX DIFFERENCES IN THE TOY-PREFERENCES OF YOUNG CHILDREN . 417
By Harold Benjamin

FOOD PREFERENCES OF THE ALBINO RAT 430
By H F. Harlow

A CLINICAL STUDY OF THE APPLICATION OF MENTAL HYGIENE TO THE TREATMENT OF CHILDREN'S SCHOOL PROBLEMS . . . 439
By John Levy

SHORT ARTICLES AND NOTES . 462
- Some New Bases for Interpretation of the IQ—Erratum (Beth Wellman) 462
- The Non-random Character of Initial Maze Behavior—Erratum (Wayne Dennis and R H Henneman) 463
- Learning of a Petz's Conure (R H Masure) 463
- Gifted Children as High-School Leaders (Frank H. Finch and Herbert A Carroll) . . 476
- A Note on the Learning of Elevated Mazes by Rats (W D Commins) . . 481
- Identical Sensory Elements versus Functional Equivalents in Visual Transfer Reactions (Fred McKinney) 483
- A Further Note on Size Preference of Albino Rats (Joseph G Yoshioka) 489
- Regarding the Law of Parsimony (Joseph W Nagge) 492

BOOKS . 495
- S Zuckerman, *The Social Life of Monkeys and Apes* (O L. Tinklepaugh) . . 495
- Books Received 497

THE EFFECT OF INACTIVITY PRODUCED BY COLD UPON LEARNING AND RETENTION IN THE COCKROACH, BLATELLA GERMANICA

From the Psychological Laboratories of Clark University

WALTER S HUNTER

INTRODUCTION

The suggestion for the present experiment came from the work by Jenkins and Dallenbach (3) on the effect of inactivity due to sleep upon the retention of nonsense syllables. These investigators, working by the method of recall, found that retention after 1, 2, 4, and 8 hours of sleep was significantly better than after corresponding periods of normal waking activities and furthermore that there was essentially no greater loss in retention after 8 hours than after 2 hours of sleep These findings have been partially confirmed by Dahl (1) using the recognition method. In Dahl's work retention after periods of 4 and 8 hour's sleep was better than after similar periods of normal waking activities For 1- and 2-hour intervals the reverse finding was secured Van Ormer (7), using the relearning method, also found that at 4 and 8 hours retention was better after interpolated sleep than after interpolated waking activities He found no difference between the two conditions for the 1-hour interval but a possible difference in favor of the sleep interval for the 2-hour period These studies clearly support the hypothesis, put forward by Jenkins and Dallenbach, that loss of retention is due to the activity which intervenes between learning and relearning, an hypothesis which relates the so-called phenomenon of forgetting to that of retroactive inhibition Perhaps no one has ever thought that sheer lapse of time was a factor in producing loss of retention, since lapse of time only permits various undetermined factors to play their rôles, as the passage of time may permit the weather to wear away the rocks The Jenkins-Dallenbach experiment, however, gives us some evidence of the rôle of specific factors in bringing about the disintegration of neural coordinations so that we need no longer say merely that forgetting is related to time The evidence, however, does not enable us to say that the neural impulses set up by muscular and glandular activities are the only important factors at work,

since during sleep the metabolic processes of the brain may well be different from their condition during waking.

The problem of the effect of inactivity upon retention is one well suited to the field of animal experimentation where studies of learning and retention have already been made with reasonable success The present paper reports the results of an experiment upon cockroaches where inactivity was produced by means of low temperatures Other experiments under way in this laboratory are using different animals and different forms of inactivity. It is not necessarily to be expected that all forms of inactivity will have the same effect, or lack of effect, upon retention, since the physiological accompaniments of the inactivity may well vary with the mode of producing the inactivity Inactivity due to cold need not have the same relation to retention as inactivity due to sleep. Indeed, the present experiments show very clearly that in the roach the learning and retention of a simple conditioned response are worse after a period of cold-inactivity than after a corresponding period during which the subject remains in a dark room at normal temperatures. This, it may be said, is not the result anticipated. It was thought that almost perfect retention might be secured if the subject's physiological processes could be held practically in quiescence during the period of rest interpolated between learning and relearning This expectation, however, was not fulfilled.

Experimental studies on learning by the cockroach have been made by Syzmanski (4) and Turner (5) These investigators trained the roach to avoid a dark compartment and remain in a lighted one. This was accomplished by shocking the animal whenever it entered the former compartment When the subject refused to enter the dark 10 times in succession learning was regarded as completed. [Turner (6) also found that the roach could learn an elevated maze having 6 culs-de-sac] These experiments have shown clearly that the roach is well adapted to learning experiments, a result confirmed by the present experiment which utilizes with but little modification the methods of Syzmanski and Turner.

Subjects, Apparatus, and Methods

The subjects of the present experiment, *Blatella germanica,* sometimes known as the croton bug, were captured by setting traps in frequented places These traps were glass battery jars the inside upper portions of which had been thoroughly greased Roaches can

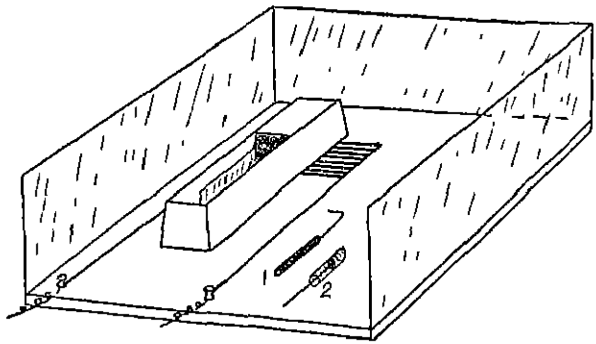


FIGURE 1

A PERSPECTIVE VIEW OF THE TRAINING APPARATUS, SHOWING THE BENT WIRE PROD, *1*, AND THE GLASS TUBE AND PISTON, *2*, USED IN TRANSPORTING THE SUBJECTS

The apparatus was walled on three sides with glass, slightly greased on the inside This enclosure made the capture of any escaped roach certain (During the course of the experiment no trained roach escaped alive to be recaught in the traps and so to reenter the experiment in the rôle of an untrained subject)

get into but not out of such receptacles For this reason the jars also served as permanent living quarters for those members of the colony who were not isolated for study Bread and water was a sufficient diet for the maintenance of the animals during the experimental period. Only roaches about 1.3 cm. long were utilized Great care was taken to insure the choice of uninjured subjects, inasmuch as animals with broken antennae or injured legs were fairly common in the colony No determination of sex was made.

The particular apparatus diagramed in Figure 1 was used only in the experiments where time intervals of 5 and 8 hours respectively were employed. Here it seemed worthwhile to improve the previously utilized apparatus by enlarging it slightly and by sloping the walls inward in order better to discourage any efforts of the roach to climb out The apparatus of Figure 1 was a 4-sided box, 9.5 ins long, 1 11-16th ins wide at the top, 2 1-16th ins wide at the bottom, and 2¾ ins high (inside measurements) The side walls were lined with glass which was coated with a very thin layer of grease (Crisco) A glass partition, 4 ins long and 1¾ ins. high, divided the illuminated portion of the box into longitudinal alleys of equal width. This partition ended 1 in from the dark chamber and 2 ins

from the other end of the apparatus A wooden top 2.5 ins. long covered and darkened one end of the apparatus Across the floor of this darkened end were strips of brass ¼ in wide, constituting the punishment grill, which were connected with an inductorium and a single dry cell. The shock used was great enough to produce a marked reaction from the roach, but it was barely supra-liminal to the dry human finger The grid was electrified only after the animal entered the dark chamber.

All illumination came from a single 200-watt blue bulb suspended 23 ins above the floor of the apparatus and so arranged that the entrance to the dark chamber was marked by a sharp shadow The temperature of the experimental room, at a point about 1 foot above the apparatus, was kept between 26° and 28° C. by means of an electric heater (The temperature of the cupboard in which the isolated animals were kept was perhaps 1° lower.) The room temperature was read at the time of each test, and the refrigerator temperature was read on a maximum-minimum thermometer whenever a roach was put in or taken out. This latter temperature ranged from 3° to 6° C

All conditions with the first experimental apparatus referred to above were the same as those just described except for the dimensions of the apparatus itself. In this case the dark chamber was 2 ins long and the illuminated portion, 6 ins long The width of the box was 1¼ ins and the height, 2 ins. (inside measurements) This apparatus was also lined with glass and fitted with the same longitudinal partition. Both apparatuses were painted inside, under the glass, with a flat black paint

In addition to the experimental (training) apparatus, there was also a practice apparatus of approximately the same size as those used for training, but lacking the longitudinal partition and the dark chamber This apparatus was used for adapting the roach to various aspects of the experimental situation prior to the beginning of regular training

The method of procedure in all experiments was as follows· (1) At a given hour on the first day a roach was taken from the colony in a small glass phial, carefully examined for physical injuries, and then placed in a small glass jar This jar was numbered and always contained a small piece of bread soaked with water. The jar and the roach were now placed in a dark cupboard (2) On the second day at the same hour the roach was placed in the practice apparatus for 10

mins and was then returned to its jar and to the cupboard. During the practice period the subject was stimulated from time to time with the bent wire prod, was tested for the strength of the shock, and in general became adapted to the experimental situation. (It will be recalled that there was no dark portion of the practice apparatus.) (3) On the third day at the same hour the roach was placed in the training apparatus where it remained until learning was completed. (*a*) If the roach belonged to the normal group it was placed again in its jar and returned to the cupboard where it stayed until the time for relearning arrived. (*b*) If the subject belonged to Experimental Group 1 it was placed in a small, numbered glass phial, which was closed at the top with thin cloth, and was placed in the refrigerator. At the close of the allotted time in the cold it was again transferred to its individual jar and placed in the cupboard for recovery until the time for relearning arrived. (*c*) Roaches belonging to Experimental Group 2 received a different treatment on the third day. At the appropriate hour, without any training, each was placed in a glass phial and put in the refrigerator for the required time. After this time had elapsed the roach was returned to its jar and put in the cupboard for recovery until the hour for its original learning test arrived.

Great pains were taken to control the time intervals involved in the experiment. For example, a roach completed its learning at 9.10 A.M. Not over one minute was required to transfer it to the cold. At 1.10 P.M. it was removed from the cold and was credited with 4 hours' exposure to the low temperature. At 2.10 P.M. it was back in the apparatus ready for relearning. No variation greater than 2 mins. occurred in this schedule.

Before the training experiment began each roach had been accustomed to the general experimental situation by means of 10 mins. spent in the practice apparatus. (The experimenter himself had become highly skilled through training more than 75 animals in preliminary work before the regular experiments began.) When the roach was placed in the experimental apparatus it usually ran at once into the darkened end where it received a shock which caused it to run back into the lighted portion. If the roach did not now of its own accord run toward the dark, it was gently touched with the bent wire prod and forced to go at least as far as the end of the glass partition. Whenever the animal went so far as this, but turned back without entering the dark, it was credited with one correct response. Fre-

quently a subject would stop at the entrance to the dark and remain quiet. If it did not move after a short pause, it was touched gently, whereupon it either entered the dark or returned to the other end of the box. Mastery of the problem was set at 9 correct responses in 10 successive trials. (The total trials required for learning, as given in the section on results, do not include these last 10 trials. However, where the total shocks for learning are recorded, all shocks received are included This results uniformly in counting one shock after the criterion of learning has been attained)

Great care was taken not to excite the roach in transferring it to or from the apparatus As an aid to this end the roach was transferred in a glass tube which contained a loosely fitting plunger. The movement of the plunger brought the animal quietly out of the tube and either into its jar or into the apparatus. To pick up the subject the tube was brought near or over the animal, who usually ran promptly inside

The following experimental program was carried out:

I. Learning followed by relearning after 1 hr.; all temperatures normal; 30 subjects

II Learning followed by relearning after 3 hrs

- (*a*) Control Group, normal temperatures, 40 subjects
- (*b*) Experimental Group 1 Learning at normal temperature, 2 hrs in cold, 1 hr recovery at normal temperature; relearning at normal temperature, 40 subjects.
- (*c*) Experimentl Group 2 Two hours in cold; 1 hr. recovery at normal temperature, learning at normal temperature; 40 subjects

III. Learning followed by relearning after 5 hrs

The same subgroups as in Experiment II, one hour being given in recovery, 40 subjects in each group.

IV Learning followed by relearning after 8 hrs.

The same subgroups as before. The two experimental groups were each given 4 hrs of cold and 4 hrs for recovery; 40 subjects in each of the three groups.

Experiments I, II, III, and IV were carried out in that order However, the three subgroups which occurred in all but the first experiment were tested in an essentially simultaneous manner. First, a subject for the control group was tested, then one for Experimental Group 1, then one for Experimental Group 2. Again a sub-

ject for the control group was chosen, followed in order by a subject for each of the experimental groups, until the total required number of subjects had been utilized. As a result of this procedure the three subgroups were equally affected by any factors which might have varied progressively as the experiment continued.

Results

Reliability. The reliabilities of the shock and trial scores were sufficiently high for the purposes of this experiment. After the experimenter had become skilled in the use of his methods of work, and before the main investigation was begun, a determination was made of the Pearson correlations between learning and relearning scores for 25 roaches. Twelve subjects were retrained 5 days after the completion of the initial learning, 8 subjects, after an interval of 3 days, and 5 subjects, after an interval of 1 day. All temperatures were normal. The correlation between the total number of trials for learning and relearning was .85±.05. For total shocks the correlation was .85±.03.

The reliabilities secured with other groups of roaches after intervals of 1, 3, 5, and 8 hrs. spent in rest at normal temperatures were as follows:

1 hour, 30 subjects, shocks, $r=.49\pm.09$; trials, $r=.40\pm.11$;
3 hours; 40 subjects; shocks, $r=.38\pm.09$; trials, $r=.30\pm.09$,
5 hours, 40 subjects; shocks, $r=.63\pm.06$, trials, $r=.31\pm.09$,
8 hours, 40 subjects, shocks, $r=.36\pm.14$, trials, $r=.11\pm.15$

Learning Followed by Relearning after One Hour. Thirty roaches were trained to avoid the dark chamber. They were then returned to the cupboard and tested for retention after 1 hour. The results secured are summarized in Table 1. They indicate a nearly perfect retention over a period of 1 hour at normal temperature.

TABLE 1
LEARNING AND RETENTION AFTER ONE HOUR

	Learning	Relearning
Av. total shocks	7±.36	1.6±.13
Range	3—14	0—4
Sigma	2.9	1.11
Av. total trials	9.2±.65	.9±.18
Range	2—20	0—7
Sigma	5.38	1.51

TABLE 2
LEARNING AND RETENTION AFTER THREE HOURS

	Group N	Group 1_3	Group i_3
	Learning		
Av. total shocks	7.7±.44	7.5±.46	10.05±3.56
Range	3—21	2—23	2—23
Sigma	4.29	4.38	5.38
Av. total trials	11.22±.87	11.32±1.01	14.52±.74
Range	2—32	1—45	2—32
Sigma	8.26	9.63	7.05
	Relearning		
Av. total shocks	3.43±.26	3.3±.26	
Range	0—10	0—15	
Sigma	2.45	2.52	
Av. total trials	4.48±.59	3.83±.39	
Range	0—21	0—20	
Sigma	5.63	3.75	

Learning Followed by Relearning after Three Hours. One hundred and twenty roaches, divided into three groups, were trained under the conditions indicated above: one group (N) learning, resting, and relearning under normal temperature conditions, one group (1_3) learning and relearning under the standard conditions, but initial learning preceded by 2 hrs. in the cold and 1 hr. in normal temperature for recovery, and the remaining group (i_3) having its initial learning preceded by 2 hrs. in the cold and 1 hr. in normal temperature for recovery.

The results are summarized in Table 2. Groups N and 1_3, who learned under identical conditions, required the same average number of total trials and shocks. Group i_3, which had been exposed to cold prior to learning, learned more slowly. The difference between the mean shocks for groups N and i_3 is 2.32±.11. The difference between the mean trials for the same groups is 3.30±.21. Both differences are significant and indicate that the preliminary exposure to cold has slowed up learning to an extent not counterbalanced by the 1 hr. of recovery.

A comparison of Groups N and 1_3 on relearning indicates that Group 1_3 has been uninfluenced by its exposure to cold during the rest period, or that some factor has counterbalanced the bad effects of the cold. There is a large and significant difference between the learning scores of Group i_3 and the relearning scores of Group 1_3. Both in trials and shocks the relearning of Group 1_3 is superior to

the learning of Group 1_3 On the basis of these results it would therefore seem reasonable to conclude that the exposure to cold has no effect upon retention but does have an injurious effect upon original learning. Later results from other groups, however, show that such a general conclusion is unjustified.

The suggestion might be made that the loss of retention by Group 1_3 occurred during the 1-hr period of recovery at normal temperature and that no loss probably occurred as a result of the 2 hrs' inactivity in the cold. This hypothesis, however, is definitely negated by results yet to be reported and also by the preceding data secured on normal retention after 1 hr which showed that retention was essentially perfect for that period of time

Learning Followed by Relearning after Five Hours In this and the following experiment the larger discrimination box, described in the section on apparatus, was used with 120 roaches, divided as usual into 3 groups Although the preceding experiment had suggested that a period of 1 hr was insufficient for complete recovery from a 2-hr exposure to cold, the present experiment again utilized the 1-hr. recovery period. The 1_5 group had 4 hrs of exposure to cold and 1 hr for recovery prior to initial learning. The 1_5 group received this same treatment between learning and relearning. Group N rested at normal temperature for 5 hrs between learning and relearning

The results on learning and retention are summarized in Table 3.

TABLE 3
LEARNING AND RETENTION AFTER FIVE HOURS

	Group N	Group 1_6	Group 1_5
	Learning		
Av total shocks	9 12± 54	8 05± 55	12 55± 82
Range	4—29	2—18	3—41
Sigma	5 16	3 5	7.79
Av total trials	14 9±1.2	13 17± 68	19 35±1 7
Range	3—46	1—30	2—53
Sigma	11 31	6 41	12 71
	Relearning		
Av total shocks	7 15± 51	11 2± 93	
Range	1—30	0—42	
Sigma	4 83	8 85	
Av total trials	11 22± 92	17 1±1 19	
Range	0—47	0—52	
Sigma	8 68	11 15	

Again we find that Groups N and 1_5 are essentially on a par in learning This is even more exactly the case if one extreme subject is thrown out of Group N. Group i_5 is again the slowest in learning. The difference between the mean shocks required in learning by Groups 1_5 and i_5 is 4.5±.16 The difference between the mean trials for the same groups is 6.18±.28. A preliminary 4-hr. exposure to cold with 1 hr for recovery apparently retards learning The data of Table 3 also indicate clearly the injurious effect of the interpolated cold upon the retention of the conditioned response in Group 1_5 Not only is the retention of this group significantly worse than the retention of Group N, but a significantly larger number of shocks and trials are required for relearning by Group 1_5 than for learning Exposure to the cold has not only cancelled all of the effects of the previous learning, but it has placed the animals of this group in a condition worse than they were in before any training had been g[iven]. The difference between the mean trials required for rele[ar]ning by Groups N and 1_5 is 5.88±.24; for mean shocks, 4.05±.1[7]. The difference between the mean shocks for the learning and relearning of Group 1_5 is 3.15±.136, for mean trials, 3.93±.156

Learning Followed by Relearning after Eight Hours Inasmuch as both the 1_5 and the i_5 groups were adversely affected by 4 hrs. in the cold and only 1 hr for recovery, it seemed quite possible that the recovery period was too short This might well be the case in spite of the fact that careful observation of the roach's behavior revealed

TABLE 4
LEARNING AND RETENTION AFTER EIGHT HOURS

	Group N	Group 1_8	Group i_8
	Learning		
Av total shocks	10.15±.49	10.68±.62	13.42±.88
Range	4—21	4—38	5—45
Sigma	4.62	5.87	8.43
Av total trials	18.1±1.15	17.8±1.15	20.25±1.45
Range	3—44	4—61	4—66
Sigma	10.67	10.9	13.78
	Relearning		
Av total shocks	7.88±.33	12.83±.79	
Range	3—16	2—33	
Sigma	3.13	7.66	
Av. total trials	14.7±.92	21.5±1.46	
Range	2—39	1—59	
Sigma	8.77	13.8	

FIGURE 2

PERCENTILE CURVES SHOWING THE EFFECTS OF EXPOSURE TO COLD UPON LEARNING WHEN MEASURED IN TERMS OF TOTAL SHOCKS

N is the normal group, 190 subjects, i_3, 40 subjects, is the group exposed to cold for 2 hours with 1 hour for recovery prior to initial learning, i_5 is a comparable group exposed to cold for 4 hours with 1 hour for recovery, i_8 is another comparable group exposed to cold for 4 hours but allowed 4 hours for recovery

no after-effects of the cold After the 1-hr recovery the roaches exposed to cold could not be distinguished from those not so exposed. In order to throw light upon this problem of the recovery period, the present experiment was carried out. Again 120 untrained roaches were used, divided into the usual three groups In the present case the interpolated normal interval was 8 hrs rest For the 1_8 group

this interval was divided into 4 hrs. in the cold and 4 hrs. recovery in normal temperature. The I_8 group was in the cold 4 hrs. and then had a 4-hr. recovery period before starting to learn the discrimination.

Table 4 summarizes the results. The evidence again clearly indicates that exposure to cold retards learning and decreases retention. The difference between the average shocks required for *learning* by Groups 1_8 and I_8 is $2.74 \pm .17$; for average trials, the difference is $2.45 \pm .292$. The difference between the average shocks required for *relearning* by Groups N and 1_8 is $4.95 \pm .13$; for average trials, the difference is $6.8 \pm .27$.

Figure 2, which will be further commented upon in the summary, indicates in terms of percentiles that the 4 hrs. of recovery allotted the roaches in the present experiment were no better than the 1 hr. of recovery previously given.

Summary and Conclusions

The learning of a simple darkness-avoiding response by the cockroach is retarded if the roach has been exposed to a continuous low temperature, 3°-6°C., for from 2 to 4 hours, and if the period for recovery is 1 hour at normal temperature. The retardation of learning is greater after 4 hours' exposure to cold than after 2 hours' exposure. An allowance of 4 hours for recovery in place of 1 hour does not better the learning performance where the subject has been exposed to the cold for 4 hours. These results are graphically summarized in the percentile curves (based on shocks) of Figure 2 where, for comparison, a curve for subjects not previously exposed to cold is included. (Percentile curves for trials show the same relationships.)

The retention of the darkness-avoiding response by roaches living in normal temperatures decreased with an increase of from 1 to 8 hours in the elapsed time. Two hours of interpolated cold and 1 hour for recovery at normal temperature was neither better nor worse for retention than 3 hours in normal temperature. Interpolated intervals of 4 hours' cold with 1 hour for recovery and 4 hours' cold with 4 hours for recovery adversely affected retention, but in equal amounts.

Inactivity produced by cold is thus detrimental, and not advantageous, to retention of the darkness-avoiding response in the cockroach. The experimental data offer no explanation for this fact. Perhaps histological studies would reveal that definite structural

changes had resulted from the low temperatures, changes which persisted at least 4 hours after the shift from low to high temperatures. Fries (2) has investigated the effect of temperature on the frequency of heart beat in the roach, *Blatta orientalis* L., and has found that the frequency of heart beat increases with temperature according to the Arrhenius equation. There is, however, a critical temperature around 10°C below which the temperature characteristic of the equation becomes higher. The degree of cold in the present experiment, ranging from 3° to 6° C., was well below such a critical temperature; but what physiological changes were thereby produced, other than a general slowing down of organic activities, cannot be said.

References

1. Dahl, A. Über den Einfluss des Schlafens auf das Wiedererkennen. *Psychol. Forsch.*, 1928, **11**, 290-301.
2. Fries, E. F. B. Temperature and frequency of heart beat in the cockroach. *J. Gen. Physiol.*, 1926, **10**, 227-237.
3. Jenkins, J. G., & Dallenbach, K. M. Obliviscence during sleep and waking. *Amer. J. Psychol.*, 1924, **35**, 605-612.
4. Syzmanski, J. S. Modification of the innate behavior of cockroaches. *J. Animal Behav.*, 1912, **2**, 81-90.
5. Turner, C. H. An experimental investigation of an apparent reversal of response to light of the roach (*Periplaneta orientalis* L.). *Biol. Bull.*, 1912, **23**, 371-386.
6. ———— Behavior of the common roach (*Periplaneta orientalis* L.) in an open maze. *Biol. Bull.*, 1913, **25**, 348-365.
7. Van Ormer, E. B. Retention after intervals of sleep and of waking. *Arch. Psychol.*, 1932, No. 137. Pp. 49.

Clark University
Worcester, Massachusetts

L'EFFET DE L'INACTIVITÉ PRODUITE PAR LE FROID SUR L'APPRENTISSAGE ET LA RÉTENTION CHEZ LA BLATTE, BLATELLA GERMANICA

(Résumé)

L'apprentissage d'une simple réponse par la blatte pour éviter l'obscurité est retardé si l'on a exposé la blatte à une basse température continue, entre 3° et 6° C., pendant 2 à 4 heures et si la durée du rétablissement est une heure à une température normale. La retardation de l'apprentissage est plus grande après que l'on a exposé la blatte au froid pendant 4 heures que pendant 2 heures. Si l'on met 4 heures pour le rétablissement au lieu d'une heure après 4 heures de froid, l'apprentissage du sujet ne devient pas meilleur. Ces résultats sont graphiquement résumés dans les courbes percentiles (basées sur les chocs) de la Figure 2, où, pour la comparaison, on

inclut une courbe pour les sujets non exposes auparavant au froid. (Les courbes percentiles pour les épreuves montrent les mêmes rapports)

La rétention de la réponse pour eviter l'obscurité chez des blattes demeurant à des températures normales a diminue avec une augmentation d'une heure à 8 heures du temps passe Deux heures de froid interpolées et une heure pour le réétablissement à une température normale n'ont été ni meilleures ni pires pour la rétention que 3 heures à une température normale Des intervalles interpolés de 4 heures de froid avec une heure pour le réétablissement et ceux de 4 heures de froid avec 4 heures pour le réétablissement ont défavorablement influencé la rétention, mais en quantités égales

L'inactivité produite par le froid est donc nuisible, et non avantageuse, à la rétention de la réponse pour éviter l'obscurité chez la blatte Les données expérimentales n'offrent aucune explication de ce fait Peut-être des études histologiques montreraient-elles que des changements définis de structure avaient résulté des basses températures, des changements qui ont persisté du moins 4 heures après que l'on a substitué une temperature élevée à une température basse HUNTER

DIE EINWIRKUNG DER DURCH KALTE VERURSACHTE UNTATIGKEIT AUF LERNEN UND GEDACHTNIS DER SCHABE, *BLATELLA GERMANICA*

(Referat)

Sezt man sie einer fortwahrenden niegrigen Temperatur von 3 bis 6° C. 2 bis 4 Stunden aus und ist die notige Erholungsperiode bei normaler Temperatur 1 Stunde, so wird die Schabe bei der Aneignung der einfachen Reaktion der Vermeidung der Dunkelheit aufgehalten Die Verzogerung des Lernens is nach 4 Studen der Kalteaussetzung grosser, als nach zweistundiger Aussetzung. Gewährt man statt einer Stunde 4 Stunden zur Erholung, so wird die Lernleistung nicht verbessert, wenn das Versuchstier der Kalte vier Stunden lang ausgesetzt worden ist Diese Befunde werden in den auf Schocke (shocks) basierten Prozentkurven (percentile curves) der Figur 2 zusammengefasst In dieser Figur ist zur Vergleichung eine Kurve von Versuchstieren, die vorher nicht der Kalte ausgesetzt worden sind, mitgegeben (Die Prozentkurven fur die Versuche erweisen die selben Verhaltnisse)

Die Beibehaltung (retention) der Reaktion der Dunkelheitsvermeidung bei Schaben, die bei normalen Temperaturen leben, nahm mit einer Zunahme von 1 bis 8 Stunden in der verstrichenen Zeit ab. Zwei Stunden interpolierter Kalte und eine Stunde zur Erholung bei normaler Temperatur ubte weder eine bessere noch eine schlimmere Wirkung auf die Beibehaltung aus, als drei Stunden bei normaler Temperatur Eingeschobene vierstunden-dauerende Kalteintervalle mit einer Stunde zur Erholung, und 4 Stunden Kälte mit 4 Stunden zur Erholung ubten eine schadliche Wirkung auf die Beibehaltung aus, aber in gleichem Masse

Die durch Kalte erzeugte Untatigkeit ist also der Beibehaltung der Reaktion der Dunkelheitsvermeidung bei der Schabe nachteilig und nicht gunstig. Die experimentellen Befunde liefern fur diese Tatsache kine Erklarung Histologische Untersuchungen wurden vielleicht offenbaren, dass durch die niedrigen Temperaturen bestimmte Strukturveranderungen erzeugt worden waren, und dass diese Veranderungen wenigstens noch 4 Stunden nach der Anderung der niedrige in hohe Temperaturen beharrten.

HUNTER

THE PERSONAL EQUATION IN RATINGS: I AN EXPERIMENTAL DETERMINATION*

From the Institute of Child Welfare of the University of California

HERBERT S. CONRAD

I INTRODUCTION

The "personal equation" between judges has long been one of the favorite criticisms of subjective ratings So well-recognized and accepted has this criticism of ratings become, that proof or measurement of the influence of the personal equation frequently seems to be considered superfluous The Army psychologists, for example, were at some pains to eliminate the personal equation from ratings of intelligence (11, pp. 438-441); yet they never undertook to evaluate the extent of the defect they were removing. An empirical determination made by the writer from the original Army data showed that the influence of the personal equation was small—in fact, practically *nil* (1). This unexpected finding has prompted us to a more extensive experimental investigation, which aims to determine the effect of the personal equation in the ratings of several judges under conditions designed to raise this effect to its *maximum* value

The "personal equation" is one of those old, traditional, and convenient terms whose meaning has tended to become somewhat indefinite with time. In the present paper, the personal equation refers to spurious differences between the means or the standard deviations of two or more judges' ratings. The *influence* of the personal equation is defined as the influence of these spurious differences upon the correlation between the ratings and an adequate criterion.

Although much used, the "personal equation" has seldom been defined except by implication, Sometimes it would appear that all deficiencies of ratings are loosely included within the term The definition above is believed to be sufficiently orthodox, tracing from original usage in astronomy, it accords with the definition implied by the Army psychologists in their careful preliminary work on the Army Alpha intelligence test (11, pp 438-439)

*Recommended for publication by Harold E Jones, accepted by Carl Murchison of the Editorial Board, and received in the Editorial Office, November 5, 1931

The very phrase, "personal *equation*," suggests that the term refers strictly only to such differences as can be eliminated by writing an equation between the respective judgments or observations of different judges An equation can, of course, be completely successful only in case there prevails some *uniform* difference in the judgments or observations of the respective persons, that is to say, η must, ideally, equal 1 00 The classic illustration of the personal equation from astronomy is the (presumably constant) difference in reaction-time of two practiced observers, who record the time of transit of a star across a hair line[1]

If judges disagree in their *relative* (i e, rank-order) placement of individual subjects—whether due to incompetence, to incomplete acquaintance with the subjects, or to prejudice and personal bias—then the differences between the judges' placements cannot be eliminated directly by an equation, and for this reason the differences should not, strictly, be described as arising (exclusively) from the "personal equation" between the judges' ratings The most that can be done in the case of relative disagreement is to attempt to discover the systematic factors (if any) which are causing the apparently chance discrepancies between the judges; or, failing this, to state the correlation ratio between the ratings of the judges, and to write the two regression equations These regression equations, however, are quite different from the personal equation As a matter of fact, unless η=1 00, they are not true equations at all for y is not a single-valued (nor, for that matter, a double-valued or triple-valued) function of x, and the "equation" used to predict the average y-value for a given x cannot be used to predict the average x-value for a given y

For purposes of clarification, it may be useful to suggest a case where the apparently chance discrepancies between the ratings of two judges are due not to any unsystematic factors, but to a personal equation, and to one other systematic factor The personal equation consists (let us say) in the fact that Judge X, on the average, rates his subjects higher by 1 than Judge Y The other systematic factor consists of a personal bias in X, whereby he over-rates blondes by 1, and under-rates brunettes by 2 As the ratings stand, no personal equation can be written which will remove all the systematic differences between the distribution of X's ratings, and the distribution of Y's ratings. Hypothetically, however, we might eliminate the influence of X's personal bias, by reducing his ratings of blondes by 1, and raising his ratings of brunettes by 2 In this way we would (in effect) cause X's ratings of both blondes and brunettes to be made on the same basis, that is to say, we would eliminate the *special* influence which the individual phenomena have upon X, and, in a sense, make X's ratings consistent with *himself* An adjustment of this kind cannot (in our under-

[1]The reader interested in a detailed technical discussion and treatment of the personal equation is referred to the study of Rhodes (7).

standing) be considered in the nature of a personal equation between Judge X and Judge Y. When, now, the personal *bias* in Judge X's ratings has been removed, the only difference remaining between the ratings of Judge X and Judge Y can be readily eliminated by the personal *equation*, $x=y+1$.

It is, perhaps, superfluous to stress that in the present paper we shall adhere strictly to the definition of the "personal equation" given above. The influence of the personal equation in ratings will be measured in terms of correlation.

II. The Influence of the Personal Equation in Terms of Correlation

Let us assume a correlation plot giving the relation between Judge A's ratings and an adequate criterion—the correlation is .65; assume another plot, giving the relation between Judge B's ratings and the same criterion—the correlation here, too, is .65; now suppose that we include *A's and B's* ratings in the *same* correlation chart, making no adjustment whatever for the personal equation between A's and B's ratings—the correlation for the combined chart becomes (say) .60. In this case, then, the disregard of the personal equation, or the personal-equation effect, is equivalent to a drop in correlation with the criterion from .65 to .60. This same personal-equation effect could have been found by determining the difference between (1) the correlation in the combined correlation chart without adjustments to A's or B's ratings (= .60), and (2) the correlation in the combined correlation chart after spurious differences in the means and SD's of the two judges' ratings have been eliminated (= .65). In the calculations of the subsequent sections, this second method is the one that has been employed.

III. The Maximum Effect of the Personal Equation, in Terms of Correlation

Our present aim is to determine the effect of the personal equation between judges' ratings, under conditions designed to raise this effect to a maximum. The personal-equation effect between two or more judges may be expected to approach its maximum under the following conditions:

1. *Single ratings of individual judges should be used.* Otherwise the effect of the personal equation is likely to be reduced, on the average, by the cancellation of errors.

2. *The judges should differ (as much as they are ever likely to*

in a practical rating situation) in their general background, and in their familiarity with statistical concepts Thus, two judges who customarily employ different levels of intelligence as reference points, or who have observed different ranges of human intelligence, are less likely to agree closely with each other in their absolute ratings Similarly, two judges unequally familiar with such concepts as mean and median, and unequally familiar with the frequency distribution which most commonly expresses individual differences in talent, are less likely to yield ratings which follow closely analogous curves

3. *The judges should each rate an EQUAL number of subjects.* Obviously, if one judge rates 100 subjects, and another rates only one, the personal equation between the judges cannot be expected to exert any highly potent statistical influence

4. *The judges should each rate a SMALL number of subjects.* In this way the true differences between judges in their standards of rating are augmented by fluctuations due to chance.

The use of a small sample may, however, involve certain dangers. There is, for example, the possibility that the particular sample may consist of subjects about whom all the judges are in exceptional agreement Such exceptional agreement may be avoided by increasing the number of judges (we have used 7), and the number of small samples (we have used 2 samples, each consisting of 12 children) Secondly, there is the statistical hazard in a small sample that chance fluctuation may occasionally diminish (though on the average it enhances) the personal equation between a pair of judges. The remedy here is to increase the number of judge-pairs. Our seven judges constitute 21 pairs—a number great enough to assure us that the *average* personal-equation effect for all these pairs will be higher in small samples of subjects than in large

It may be pointed out that, since judges are not infrequently called upon to rate small groups [as in the investigation of Webb (10)], the deliberate use of a small sample in the present study is supported by experimental as well as by statistical considerations

5. *The judges should rate the subjects without consultation,* either about the rating properly to be assigned in any particular case, or about the exact meaning of the various divisions of the rating scale (or other rating instrument employed) In a study of the *maximum* extent of the personal equation, the only legitimate source of uniformity of procedure or interpretation lies in the experimenter's directions, and in the agreement intrinsic to the judges.

6 *Raw, unadjusted ratings should be used* Almost any manipu-

lation (such as conversion to sigma-scores) is more likely to reduce than to augment the spurious differences between judges

7 *The ratings of each judge should correlate high with the true (or best available) criterion* (2) For if each individual judge's ratings correlated *zero* with the true measures, then the ratings would have no relation to the true measures in the first place; in short, the ratings would themselves be altogether invalid or spurious; and hence elimination of spurious differences between the individual judges in means and standard deviations could result in no meaningful improvement whatsoever.

The treatment below generally assumes that, for each personal-equation effect, two judges only are involved It is clear that when ratings have been made by more than two judges, determining the personal-equation effect between all the possible pairs of judges enables a very full use of the experimental material However, it is also possible to determine the correlation between a criterion and the ratings of more than two judges, both with and without adjustments for the personal equations among them.[2] The use of pairs of judges in the present study is, then, inspired by expedience rather than necessity The use of single pairs of judges involves one definite advantage; viz, that a negative personal-equation effect between one pair of judges is not counterbalanced by a positive effect between another pair By the technique employed, every personal-equation effect between every pair of judges is fully revealed

IV The Sample and the Data

For the experimental study of the personal equation, we have available ratings of the intelligence of 14 children by three teachers in the Children's Community (a cooperative nursery school in Berkeley, California), and ratings of the intelligence of 31 children by five teachers in the Nursery School of the Institute of Child Welfare The rating scale on which judgments were recorded is as follows:[3]

RATING SCALE FOR PRESCHOOL CHILDREN

Rate each child on a 5-point scale, 5 indicating a very high degree of the trait, 3 an average degree of the trait, and 1 a very low degree of the trait

---

[2]See footnote 7

[3]For the collection of the original data used in the present report, the writer is indebted to Dr Harold E Jones, Director of Research at the Institute of Child Welfare, and to Dr A S Jaffa Dr Jaffa administered the mental tests which have been employed as the criterion of the ratings collected by Dr. Jones

Do not overestimate In an unselected sample of children from homes of the type which supply our nursery school membership, about 35% would normally be rated 3, about 20% would be rated 2 or 4, and about 10% would be rated 1 or 5

In making your rating, take the child's chronological age into consideration Thus, a 30-months-old child should be rated with reference to others of a similar age, and not considered in comparison with older or younger children

Try to avoid "halo" effects Because you have rated a child high in one trait, do not feel that you have to rate him high in every other

Child	Age	Mental alertness	Ability to learn	Play initiative	Speed of comprehension	Language organization
Andrew M	1-8					
Sheila H.	2-0					
Ross J	2-2					
.	.					

Each child's rating in "intelligence" was considered to be his average rating in the five traits weighted as follows mental alertness, 1; ability to learn, 1; play initiative, ½, speed of comprehension, 1, and language organization, 2

We shall name the three teachers who made ratings in the Children's Community, A, B, and C; the five teachers at the Institute Nursery School are named D, E, F, G, and H From the 31 children rated at the Institute School, 12 children who had been rated by all five judges were selected at random. Similarly, from the 14 children rated at the Children's Community, 12 children who had been rated by both Judge A and Judge B were selected (Judge C did not rate 12 children in common with either of the other judges) These two sets of 12 children from different schools composed the small, equal samples rated by judges of different background and statistical training, which were specified in the preceding section as necessary for a study of the maximum extent of the personal equation

The extent to which the seven judges differ in "background and statistical training" can hardly be stated in exact terms. The judges in the Institute school are all professional nursery-school teachers, who have had at least as much academic training as is represented by the M A degree. This academic training, however, lies in various fields (psychology, child development, hygiene, and home economics), and was received by the respective teachers at different times, and in different places The years of professional

experience of the Institute teachers with preschool children vary from 6 to 30 months. At the Children's Community, the two judges are mothers who have served as teachers in the school, but who have relatively little special experience or academic preparation in child development or in psychological measurement. It appears safe to say that among the five judges at the Institute school, the differences in background and statistical training are—at the minimum—appreciable, and that the differences between the Institute judges on the one hand, and the Children's Community judges on the other, are of a fairly marked degree.

As an adequate criterion of intelligence against which the individual ratings might be checked, the average of from 5 to 9 IQ's in a mental-test battery of standard intelligence tests[4] was used. Table 1 gives the age, the average IQ, and the ratings of each child used in the present study.

V. The Magnitude of the Personal Equation Among Judges in the Same Nursery School

The correlation between the criterion of intelligence and each judge's ratings has already been given in the bottom row of Table 1. The five judges in the Institute school constitute 10 possible pairs, the two judges in the Children's Community constitute a single pair, making a total of 11 pairs. Between the members of each of these 11 pairs, a "personal equation" exists. We proceed at once to determine the magnitude of this personal equation, in terms of correlation.

The ratings of Judge A correlate .633 with the criterion ($n=12$),[5] the ratings of Judge B correlate .793 ($n=12$).[5] If no personal equation existed between these two judges, and a combined contingency table were drawn up, the correlation would be[6] simply the average of .633 and .793, or .713 ($n=24$).[5] Actually, however, the

---

[4]Viz., the Merrill-Palmer Pre-School Scale, the Gesell battery (IQ obtained from an arrangement of the items into an age-scale by G. A. Jaederholm), the Minnesota Preschool Test (a preliminary form kindly made available for our use by Dr. Goodenough), and the Stanford-Binet. *All* of the children were given (at suitable intervals over a period of 1.2 years) the Merrill-Palmer, Gesell, and Minnesota tests, and one retest on the Minnesota, the remainder of the tests administered were either the Kuhlmann-Binet, or retests. The average number of tests administered to each child is 6.7, the median number is 7.

[5]The reader is asked to remember that these small populations are deliberate, for the purpose of determining the *maximum* personal equation (see Section III).

[6]According to the formula in reference (11), p. 439. Cf. footnote 12.

TABLE 1

ORIGINAL DATA

Institute School								Children's Community				
			Intelligence rating by judge								Intelligence rating by judge	
Child	Age in months	Average IQ	D	E	F	G	H	Child	Age in months	Average IQ	A	B
1	38	120	3.9	3 3	3 8	3 1	3 7	13	39	87	1 4	2 1
2	40	100	2 2	2 1	1.5	1.6	2 1	14	43	138	4 9	4 8
3	41	109	4 3	3 4	1 8	2.9	3 4	15	46	125	5 0	4 6
4	42	117	3.5	3 5	2.4	3 0	3 2	16	47	103	4.1	3 4
5	44	118	3 4	2 4	2.4	3.0	2 4	17	57	115	4.4	3 5
6	45	133	4 3	4 2	3 9	4 9	4 1	18	57	118	4 5	4 6
7	46	99	2 0	2 3	1 9	2 6	2 6	19	57	119	4 5	4 1
8	50	128	4 1	4 1	4 3	4 9	4.8	20	60	125	2 7	3 2
9	51	102	3 6	3 3	2 6	3 3	3 0	21	61	123	3 0	4 8
10	53	121	3 5	3 0	3 6	3 9	4 3	22	33	114	3 8	3 4
11	53	102	2 3	2.7	2.3	2 9	3 5	23	38	116	3 0	4 5
12	54	140	3 6	3 8	4 5	2 9	4 1	24	67	143	4 7	4 8
Mean.		116 25	3.392	3 175	2 917	3.250	3 433			119 33	3 833	3 983
σ		13 039	7675	6622	9974	8912	7835			14 011	1 0459	8224
r,* rating x IQ			669	746	881	613	715				633	793

*All correlations reported in this paper are Pearson r's, and without exception have been computed from the raw scores without grouping [reference (4), formula 130]

TABLE 2
MAGNITUDE OF THE PERSONAL EQUATION

Judge pairs	Correlation in combined contingency table, with personal equation present	Correlation in combined contingency table, with personal equation eliminated	Difference between (3) and (4) personal equation
(1)	(2)	(3)	(4)
A-B	.696	.713	017
D-E	695	708	013
D-F	756	.775	019
D-G	635	641	006
D-H	692	692	000
E-F	802	814	012
E-G	662	680	018
E-H	716	731	015
F-G	742	.747	005
F-H	771	.798	027
G-H	.656	.664	008
Average	711	724	013

combined contingency table yields a correlation of only .696 The "personal equation" has, accordingly, deflated the ideal correlation by .017.

A similar technique applied to the other pairs of judges yields the results given in Table 2.

For the 11 pairs of judges in Table 2, both the mean and the median influence of the personal equation, under conditions calculated to magnify rather than reduce the influence, are in terms of correlation only .013 [7]

It may be well, here, to give a concrete illustration of the relative insignificance of the personal equation, in a group which is larger than our deliberately small sample of 12. Chart 1 presents each of two judges' ratings of 30 children, plotted against the average IQ's of these children. The crosses constitute a correlation plot of the

[7]If we combine the contingency tables of *all* the judges of the Institute school into a *single* table, the correlation resulting is 706, with no deflation from the personal equation, this would be 725 (i.e., the average of 669, 746, 881, 613, and 715—cf Table 1) The personal-equation effect, therefore, is 725 less 706, or 019 This is virtually the same as was derived above by combining only two contingency tables at a time, and averaging the 11 separate personal-equation effects

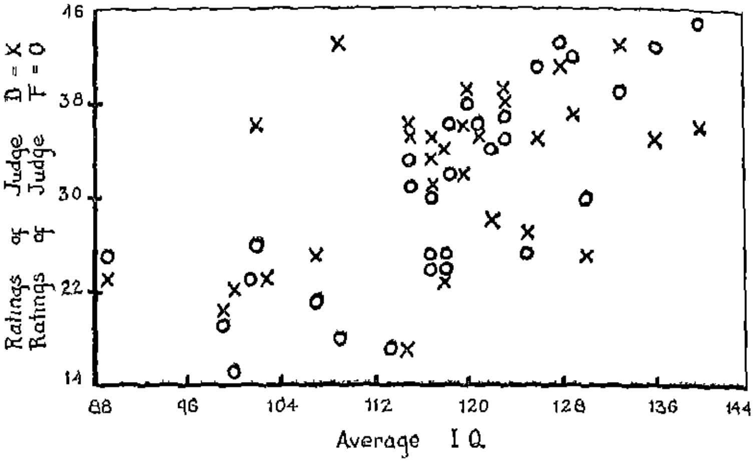


CHART 1
RATINGS OF JUDGES D AND F PLOTTED AGAINST AVERAGE IQ

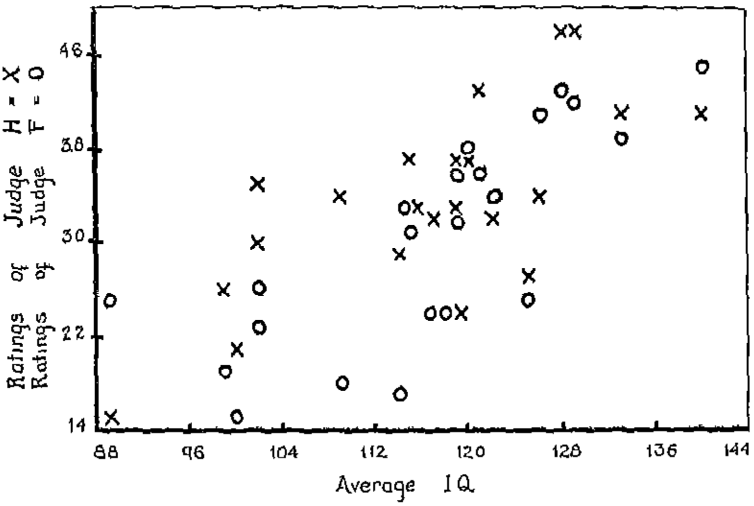


CHART 2
RATINGS OF JUDGES F AND H PLOTTED AGAINST AVERAGE IQ

ratings of Judge D, the circles constitute a correlation plot of the ratings of Judge F. The intermingling of these circles and crosses is evidence that the inclusion of the ratings of each judge into a single correlation chart has had no significant effect upon the correlation of the ratings with IQ's. In point of fact, if the means and standard deviations of the two judges' ratings had been absolutely identical—that is to say, if there were no personal equation whatever between the two judges—the correlation in the combined chart would be only .006 different from what it actually is. It may be recalled (Table 2) that for these same judges, in the deliberately smaller sample of 12 cases, the personal-equation effect is over three times as large (viz., .019): a fact which confirms the validity of the technique herein employed to discover the *maximum* effect of the personal equation (cf. Section III).

The largest personal-equation effect in Table 2 occurs upon the combination of the contingency table of Judge F and of Judge H. A correlation plot of the ratings of both these judges, for the total 22 children rated by each, is presented as Chart 2. Here also, mere inspection is sufficient to indicate that the effect of the personal equation upon the correlation of the ratings with IQ's is insignificant; by computation, the effect (in terms of r) is only .013. This, again, is less than half the personal-equation effect in the smaller sample of 12 cases (cf. Table 2).

VI. The Magnitude of the Personal Equation Among Judges in Different Nursery Schools

Mention has already been made of the fact that differences in background and statistical training are, on the whole, smaller among judges in the same school, than among judges in different schools. We might, therefore, expect the personal-equation effect between the teachers in different schools to be possibly greater than that between the teachers in the same school. To the determination of this possibility we now proceed.

The actual correlation resulting from the combination of the contingency tables of Judge A (Children's Community) and of (say) Judge H (Institute school) may readily be computed, in the usual way, as .668. The problem remaining is to eliminate the effect upon this correlation of spurious differences between the means and the standard deviations of the two judges. The mean of Judge A's rating is 3.833, of Judge H's rating, 3.433. A's mean rating is higher than H's—as it should be, since the mean

IQ of the group rated by A is 119.33, whereas that of the group rated by H is 116.25 (cf. Table 1). However, A's mean rating, while differing in the right direction from H's, differs *too much*. On the basis of Judge A's regression of IQ on rating, we determine that a difference of 3.08 points in IQ is accompanied by a difference of only .146 points of rating. Judge A's mean rating, therefore, should really be .146 higher than Judge H's—or 3.579.

Turning now to the standard deviations, we find that the SD of Judge A's ratings is 1.0459, of Judge H's ratings, .7835. Judge A's SD is higher—as it should be, since the standard deviation of the IQ's of A's group is 14.011; whereas the standard deviation of the IQ's of H's group is 13.039 (cf. Table 1). However, here again Judge A, while differing from Judge H in the right direction, differs *too much*. We assume that the ratio between the SD's of the IQ's ought to prevail also between the SD's of the ratings. Solving, then, the proportion 13.039 : 14.011 = .7835 : x, we find that the standard deviation of Judge A's ratings should be .8419. It should be specially noted that altering the mean and standard deviation of Judge A's distribution of ratings does not affect the correlation between A's ratings and the criterion of IQ's (3).

We have, now, the following situation: Judge A's ratings correlate .633 with the criterion, Judge H's, .715. The mean and SD of Judge A's ratings have been adjusted to 3.579 and .8419, respectively; this eliminates the personal equation between Judges A and H. The combined contingency table of Judge A (adjusted) and Judge H yields a correlation of .674,[8] the combined contingency table of Judge A (unadjusted) and Judge H yielded a correlation of .668. The "personal equation" between A and H, therefore, amounts to .674 *minus* .668, or but .006 points of correlation.[9]

The technique presented above has been applied to all judge-pairs the members of which come from different schools. The results are given in Table 3.

For the 10 pairs of judges in Table 3, the average personal-equation effect, under conditions calculated to magnify rather than reduce its effect, is, in terms of correlation, only .023, the median is

---

[8]Using the formula on p. 439 of reference (11) (given in footnote 12, below). In order to obtain the statistical constants of the combined distributions, required by this formula, use is made of the appropriate formulas from reference (12), p. 142.

[9]It is obvious, in the above, that instead of adjusting Judge A's ratings and leaving H's unaltered, we might have adjusted H's ratings and left A's unaltered. If H's ratings are adjusted instead of A's, the personal equation is calculated as .005—or virtually the same as it was before. This virtual equality of result, regardless of which judge's ratings are adjusted, holds not only for Judges A and H, but for every judge-pair in Table 3. Only one set of figures, therefore—based on the adjustment of the ratings of the judges in the Community school—is included in Table 3.

TABLE 3

MAGNITUDE OF THE PERSONAL EQUATION AMONG JUDGES FROM DIFFERENT SCHOOLS

Judge pairs	Correlation in combined contingency table, with personal equation present	Correlation in combined contingency table, with personal equation eliminated*	Difference between (2) and (3) personal equation
(1)	(2)	(3)	(4)
A-D	648	653	005
A-E	657	.689	032
A-F	.726	749	023
A-G	.626	626	000
A-H	668	674	006
B-D	724	738	014
B-E	.728	.774	046
B-F	772	835	063
B-G	.684	712	028
B-H	748	759	011
Average	698	.721	023

*Cf. footnote 9

between .014 and 023 The greatest personal-equation effect, 063, occurs between Judges B and F, it is worth noting that the ratings of F correlate higher than those of any other judge with the criterion, and the ratings of B correlate next highest

VII THE MAGNITUDE OF THE "TEST EQUATION" BETWEEN OBJECTIVE INTELLIGENCE TESTS

Attention may be called here to the fact that so-called objective tests have what might be termed a "test equation" That is to say, two tests applied to the *same* group yield *different* means and *different* standard deviations (6, 8)—especially if the group contains (as in our procedure above) only 12 cases. For comparative purposes, we have determined, for the same sample used in the study of the personal equation, the effect of the "test equation" between the Merrill-Palmer and a preliminary form of the Minnesota Preschool Scale As a criterion against which to check the IQ for each scale, the same average IQ from a battery of tests employed in the study of the personal equation has been used The results of this investigation will not be presented in detail Suffice it to say that when the same

children are given each test (both Merrill-Palmer and Minnesota), the test-equation effect between the two tests amounts to .012 (Institute school) and .014 (Community school). When the test equation between the Merrill-Palmer scale (administered to the children in the Institute school) and the Minnesota scale (administered to children in the Community school) is determined, the effect comes to .007, between the Minnesota (administered in the Institute school) and the Merrill-Palmer (administered in the Community school), the effect is .011 By comparison with the average personal-equation effect, these test-equation effects are certainly appreciable,[10] but neither the personal-equation nor the test-equation effects are large enough to be generally significant

VIII CRITICISM OF RESULTS. THE FACTOR OF SKEWNESS

The results of the preceding sections are perhaps as surprising to the reader as they were to the writer. Is there not some important flaw in the conditions or conduct of the experiment? And is there not—despite the "common-sense" demonstration given in Section V—some important omission in the statistical technique?

Within the knowledge of the writer, the only omission in the statistical technique is the neglect, up to this point, of the factor of skewness The literature on ratings emphasizes the difference between judges in the means and standard deviations of their ratings (13), with respect to the remaining aspect of the frequency curve which is not necessarily reflected in the mean and the standard deviation—namely skewness—the literature mentions the widespread tendency of judges to refuse to assign low ratings (5). The consequence of this tendency would be a negative skew in the frequency curve of the judge's ratings. This negative skewness has been mentioned as a more or less *constant* error, common to all judges,[11] for this reason, skewness was explicitly neglected in the statistical treatment of the preceding section

At this point, however, it may be desirable to proceed more critically, and examine in detail the possible effect of differences among the judges in skewness Suppose Judge X_1's ratings of a group of children are given in frequency curve x_1, and Judge X_2's ratings of the same children are given

[10]To the objection that it is not fair to use preschool tests of intelligence, because these are still not fully developed, one might reply in kind that neither are rating scales fully developed, nor judges adequately trained or coached for their task of rating

[11]An effort is made in recent rating scales to avoid this error by explicitly warning the judge against it, or by specifying the number of cases which the judge is expected to place in each of the divisions of the rating scale.

in curve x_2. The curves (let us suppose) are equal in means and in standard deviations, but differ in skewness. If we combine curves x_1 and x_2 into a single frequency distribution, what will be the mean and standard deviation of the combined curve? Both the mean and standard deviation will be exactly the same as for each individual curve (12, p. 142).

Suppose, now, that Judge X_1's ratings correlate .65 with an adequate criterion, and Judge X_2's the same. If we throw into a single, combined contingency table the individual contingency tables of Judge X_1 and of Judge X_2, what is the correlation in the combined table? The correlation is .65[12]—exactly the same as for each individual contingency table. In terms of correlation, then, differences in skewness may be neglected. The slightest reflection, however, will show that neglect of differences among judges in skewness—if the differences are large enough—will result in serious injustice to the children who are judged. For suppose that Judge K's ratings, on the average, equal Judge L's, but K's ratings range from -4.0σ to $+1.0\sigma$, whereas Judge L's ratings (on comparable but not identical) children run from -1.0σ to $+4.0\sigma$. Then it is clear that on the basis of standard (i.e., sigma) scores, Judge L's children will (let us say) be awarded scholarships much more often than the children of Judge K. We conclude, then, that so far as skewness is concerned, the simple correlation criterion which we have employed to measure the influence of the personal equation is inadequate.

Before proceeding to assess the importance of differences among judges in

---

[12]This follows from the formula for r_{xy}, the correlation in the combined table (11, p. 439):

$$N\sigma_x\sigma_y r_{xy} = N_1\sigma_{x_1}\sigma_{y_1}r_{x_1y_1} + N_2\sigma_{x_2}\sigma_{y_2}r_{x_2y_2} + \frac{N_1N_2}{N}(\bar{x}_1-\bar{x}_2)(\bar{y}_1-\bar{y}_2)$$

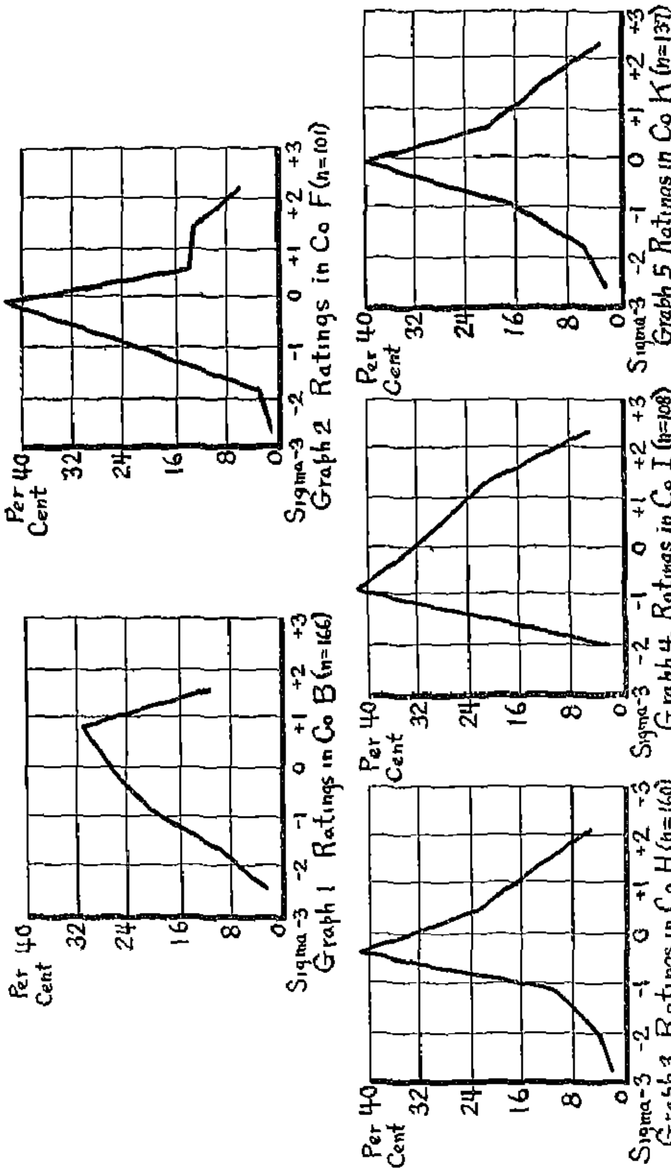
In this formula, x_1=the ratings of Judge X_1, x_2=the ratings of Judge X_2, $y_1=y_2$=the criterion scores of the group of children who were judged, x=the ratings in the combined distribution, y=the criterion scores in the combined distribution, $N=(N_1+N_2)$=the number of cases in the combined distribution. $\bar{x}_1$ is the mean of the x_1 distribution, $\bar{x}_2$ is the mean of the x_2 distribution, etc. Since in the case under discussion $\bar{y}_1=\bar{y}_2$, and $\bar{x}_1=\bar{x}_2$, the last term of the formula vanishes.

Further, since $\sigma_{x_1}=\sigma_{x_2}=\sigma_x$, and $\sigma_{y_1}=\sigma_{y_2}=\sigma_y$, and $N=2N_1=2N_2$, the formula may be re-written $2\sigma_x\sigma_y r_{xy}=\sigma_x\sigma_y r_{x_1y_1}+\sigma_x\sigma_y r_{x_2y_2}$

Whence,

$$r=\frac{\sigma_x\sigma_y(r_{x_1y_1}+r_{x_2y_2})}{2\sigma_x\sigma_y}$$

$$r=\frac{r_{x_1y_1}+r_{x_2y_2}}{2}$$


Per Cent
40
32
24
16
8
0
Sigma -3 -2 -1 0 +1 +2 +3
Graph 1 Ratings in Co B (n=166)
Graph 2 Ratings in Co F (n=101)
Graph 3 Ratings in Co H (n=160)
Graph 4 Ratings in Co I (n=108)
Graph 5 Ratings in Co K (n=137)

skewness, it may be well first to examine several frequency curves of ratings Perhaps after spurious differences in means and standard deviations have been eliminated, differences in skewness are obviously negligible. Graphs 1-5 give the frequency curves of ratings in five Companies at Camp Meade [13]

We should expect [on the basis of the curve of Army intellect yielded by the Alpha test (9, p 7)] that the distribution of ratings of intelligence in these Companies should be somewhat positively skewed A glance at the curves shows that this expectation is fulfilled in all cases except Company B To what extent Company B's intelligence really is atypically skewed, as the curve of ratings would suggest, and to what extent the ratings in the other Companies fail to reveal cases of atypical skew, it is not possible exactly to state [14] The fact that the correlation with Beta scores drops when the ratings of Companies B, H, and K are arbitrarily equalized by the use of a normal curve (1) indicates that the unusual skew in the ratings of Company B is not altogether a result of error on the part of the ratings. All in all, then, we may conclude—so far as these Army results permit a conclusion—that *illegitimate* differences in skewness are not present to an important extent.

The Army data, while suggestive, are inconclusive for lack of an adequate independent criterion against which to check the ratings. In the Institute nursery school, each of 25 children were rated by four judges The criterion of intelligence for this group, as for the two groups of 12 children used in the preceding section of this paper,[15] consists of the average

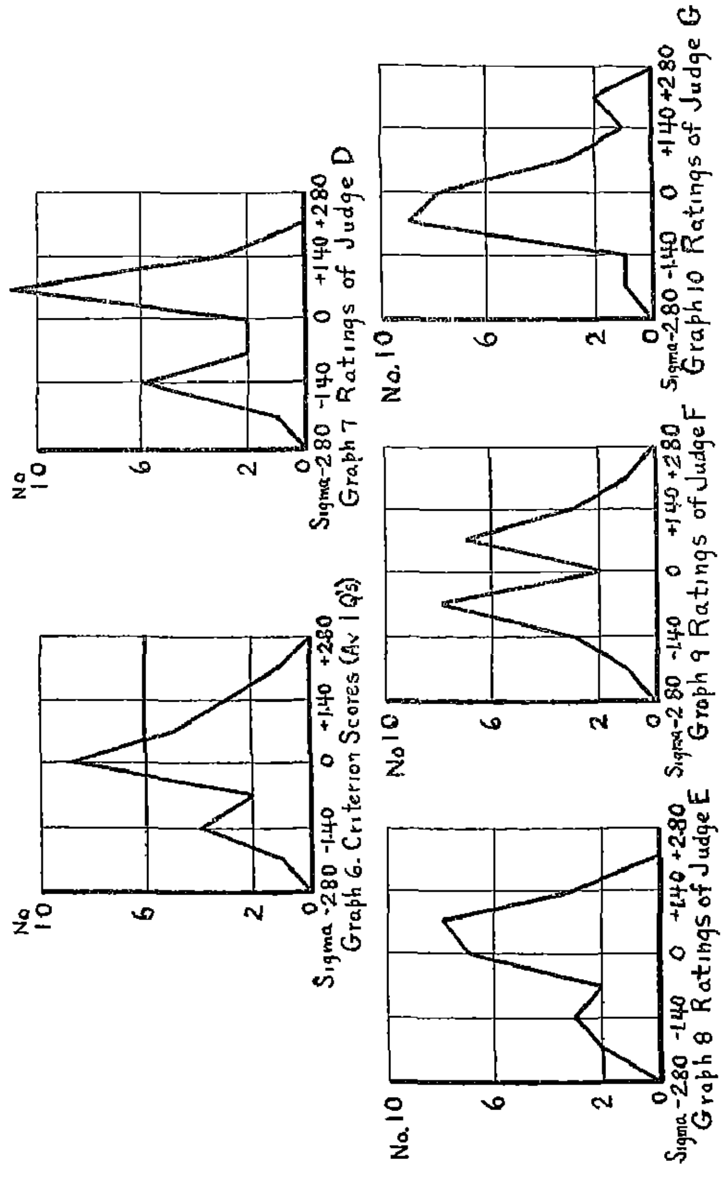
---

[13]In Graphs 1-5, the x-axis in each case represents sigma deviations from the mean of the group That is to say, the raw ratings in each Company have been translated into their sigma equivalents As a result of this translation, the means of all Companies are placed at a uniform figure (namely, 0σ), and all deviations from the mean are measured off according to a uniform or comparable unit (namely, the standard deviation of the Company) This uniformity in means and standard deviations is desirable, when skewness exclusively is under consideration The y-axis in Graphs 1-5 gives percentages instead of the original frequencies, again in order to make the frequency curves more readily comparable with respect to skewness

The original data for Graphs 1-5 are given in reference (11), Tables 110 and 118

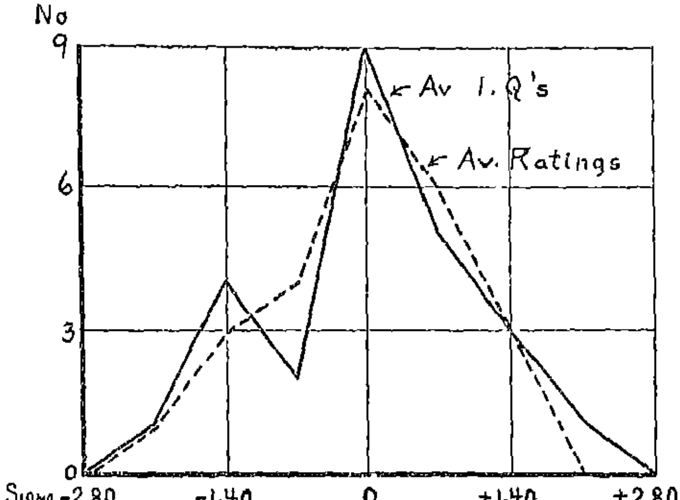
[14]Alpha and Beta scores, which might serve as criteria, are available for each Company (11, Tables 110 and 118) But the frequency curves of the Alpha scores are invalidated by an extreme piling up of cases at the low-score end, and the frequency curves of Beta scores seem to be invalidated by a spurious concentration of cases at the other end, a result of coarseness of units at the upper end, of several of the component Beta subtests—cf (11, pp 627-628) It is interesting to notice (cf Table 118) that the curve of Company B's Beta scores, like its curve of ratings, exhibits a greater negative skew than any other Company's (but Company B's ratings differ from the other Companies *more* than do its Beta scores)

[15]The group of 12 children used in the preceding section has here been enlarged to 25, because frequency curves based on only 12 cases would be too irregular to be meaningful for the present problem


Graph 6. Criterion Scores (Av I Q's)
Graph 7 Ratings of Judge D
Graph 8 Ratings of Judge E
Graph 9 Ratings of Judge F
Graph 10 Ratings of Judge G
Sigma -280 -140 0 +140 +280
No 10 6 2 0

of from 5 to 9 IQ's in a mental-test battery of standard intelligence tests.[16] Graphs 6-10 present frequency curves of the ratings of each judge, and of the criterion As in the case of the Army curves, sigma-scores have been used as the x-axis.

We should hesitate to affirm dogmatically, on the basis of Graphs 6-10, that differences in skewness between judges are or are not important It is noteworthy, however, that, if for each child an average is formed of the sigma-score ratings of the four judges, the frequency curve of the average ratings is practically identical with that of the criterion IQ's (see Graph 11). In other words, no matter how significant the factor of skewness may be in individual judges' ratings, this significance practically vanishes if several (say three or more) judges' ratings are averaged



Graph 11. Average Ratings vs Average I Q's

Where special refinement is required, or where only a single judge is available, it is suggested that ratings be adjusted on the basis of a frequency curve midway between the actual frequency curve and the theoretically correct one Arbitrary adjustment of all ratings to

[16]Mean number of tests administered to each child, 7 32, median number, 7.

the theoretically correct form (the method used by the Army psychologists) is not likely to be most satisfactory, it makes the indefensible assumption that all divergences from the theoretical curve are due solely to errors of judgment, and not at all to fluctuations or selective factors in sampling.

IX. Further Criticism of Results: The Conditions of the Experiment

The preceding section discussed a possible criticism of the statistical technique; the present section will consider possible defects in the conditions of the experiment. To what extent is the small influence of the personal equation, as found in Section V, due to (1) conference among the nursery-school teachers at the time of rating, (2) discussion or conversation among the nursery-school teachers, at various times, concerning the intellectual characteristics of the children; and (3) knowedge of the chidren's IQ's?

The first item may be promptly dismissed no conference or discussion took place among the teachers at the time of the rating. With respect to the second item, there is no question that a certain amount of discussion and conversation between teachers in the same school occasionally does take place. However, between the teachers at the Children's Community and those at the Institute, discussion concerning the respective groups of children rarely, if ever, occurs—and we have already seen that the personal-equation effect among judges in different schools is only a trifle greater than that among teachers in the same school. Moreover, this slight difference between the two personal-equation effects is readily explainable by chance, or by greater differences among the judges of different schools in background and in statistical training

It deserves recognition that, between the teachers in the same nursery school, discussion or conversation is usually about the child who is rather obviously exceptional To the extent that discussion actually occurs, and includes a consequential number of children, and actually alters teachers' evaluations, we should expect spurious agreement between judges in their *ranking* (i e , relative placing) of the children But, even so, no necessary reduction of the personal equation between the judges' *ratings* would occur, for each judge might still employ a different rating to express the mean of the group being rated; and each judge might still employ a different sigma-unit to express deviations from the mean The "personal equation,"

it must be remembered, is *not* a matter of *relative* agreement or disagreement between judges, but a question of absolute differences in the mean, standard deviation, and skewness of the distribution of each judge's ratings.

With respect to the third item, viz., knowledge of the children's IQ's,[17] the identical argument applies as in the preceding paragraph. In fact, the higher each individual judge's ratings correlate with average IQ, the *greater* is the likelihood (other things being equal) of a large personal equation (see Section III). Fortunately or unfortunately, however, the teachers rated the children not on their IQ's, but on five more or less separate, and somewhat specific traits, viz., mental alertness, ability to learn, play initiative, speed of comprehension, and language organization (see Section IV). Of course, halo-effect in the ratings probably made these traits less specific than was desired. But even if the teachers did rate the children on the basis of a single general impression, it is believed that, with the given questionnaire and under the experimental conditions, this general impression proceeded from observation of the children's behavior, rather than from any knowledge of IQ's. We may suggest, in this connection, that nursery-school teachers do not appear to take any too seriously the IQ's of young children on present preschool intelligence tests. With their own exceptional opportunities for observation, the teachers appear to consider their own judgment of the child to be more important than (or at least just as important as) the IQ. The teachers, moreover, appear to consider the IQ as representing

[17]The ratings used in the present study were made in 1929. Judges D and E, however, knew and rated the children's intelligence in 1928 as well as in 1929. After the ratings of 1928, the teachers were given some knowledge of the children's IQ's. If this knowledge were remembered, and were influential, then the ratings of Judges D and E in 1929 ought to correlate higher with the children's average IQ, than the ratings of the same judges on the same children in 1928. No such higher correlation is observed.

Judge	Rating	Correlation with average IQ, 1928	Correlation with average IQ, 1929	No. of cases
D	Average of two ratings, made each year	73	63	23
E	Single rating, made each year	79	83	19
	Average correlation	76	73	

a more or less special, typically "mental test" kind of ability. This "mental test" ability is admitted to overlap, but not to coincide with the equally significant abilities discerned in daily observation

X. Special Cases

The two preceding sections have failed to uncover any reason for rejecting the evidence adduced from the Army *Memoirs* (1), or from our own experiment Special circumstances are readily imaginable, however, under which the "personal equation" may become quite significant Thus, the judges may differ widely in the level of intelligence which they customarily employ as a reference point, or in the range of human intelligence which they have observed The judges may be very unequally familiar with statistical concepts, or with the frequency distribution which most commonly expresses individual differences in talent The judges may be given instructions which are so incomplete or non-specific that individual differences in background and statistical training are given full rein. The judges may differ seriously in their interpretations of the instructions accompanying the rating scale, or even choose (whether deliberately or not) to disregard the instructions, and rate according to their own intuition or habit Such disregard of instructions is most likely to occur when judges are rushed for time, or uninterested, or unconscientious. Finally, the judges may derive personal advantage through consistent under- or over-rating, and fail to resist this temptation uniformly

It follows, from the above, that for ideal results, not only must the rating scale and its directions be very carefully constructed, but judges must either be carefully trained in rating, or carefully selected, or both. In any case, the best results are likely to be obtained if sigma-scores of each rater are used, rather than raw scores. If the means and standard deviations of the sample rated by each judge can be more or less adjusted on the basis of independent information (such as occupational level of the fathers, social status of the parents, or scholastic record of the group), so much the better Cases arise, however, in which the group being rated is so small as to render the calculation or adjustment of mean and standard deviation impractical. In such situations, especially, an extra effort should be made to secure ratings on each subject from *several* judges, these judges, moreover, should be especially well equipped and well trained for their task If these conditions cannot be fulfilled, the ratings should preferably be replaced or be supplemented by appropriate tests of a known (not assumed) degree of validity

XI Summary and Conclusions

In their careful study of the validity of the Alpha and the Beta tests, the Army psychologists made correlations with officers' ratings. In order to eliminate the "personal equation" from the ratings, arbitrary adjustments were made on the basis of the normal curve. These adjustments, the writer has shown (1), failed to improve the correlation of the ratings with the criterion—indicating that the personal-equation effect is probably much smaller than has generally been supposed.

These results prompted a further investigation of the influence of the personal equation in a preschool sample for which fairly adequate test measures of intelligence were available (viz., the average of from 5 to 9 IQ's in standard intelligence tests). The mean and standard deviation of each judge's ratings in this sample were objectively adjusted on the basis of the mean and the standard deviation of the distribution of average IQ's. The personal-equation effect was then defined as the difference between the correlation of the *adjusted* ratings with the criterion of IQ's, and the correlation of the *unadjusted* ratings with the same criterion. In order that the personal-equation effect should be assessed at its maximum, this procedure was applied to two deliberately small groups of 12 children each, selected at random from two nursery schools. The judges in these schools differ considerably among themselves in their professional and educational background.

Under these rather special adverse conditions, the mean (also median) personal-equation effect between teachers in the same school comes to no more than .013 points of correlation; the mean effect between teachers in different schools comes to only .023 points of correlation. The range of the total 21 determinations of the influence of the personal equation is from .000 to .063; the mean effect of the personal equation is to reduce a correlation of .722 between ratings and criterion to .705—a difference of only .017.[18]

To the above findings the objection may be raised that the statistical

[18]It is not at all unusual, when relations are *in the neighborhood of .70 or below,* for elaborate statistical adjustments to fail to improve a correlation by more than .02 or .03. To give a specific instance from the *Memoirs* (11, Table 101), redistributing as many as 30% of the cases in a correlation chart (in a case where this portion of the population was quite inadequately measured by the Alpha test) failed to raise an original correlation of .5867 to more than .6138. This instance from the *Memoirs* is not at all unique.

technique employed ignores spurious differences between the judges in the skewness of their ratings Such differences of skewness, however, would in no way affect the correlation with a criterion, nevertheless, they might, if sufficiently large, result in some injustice to the children being rated The exact importance of this factor of skewness has not been quantitatively evaluated in this paper. It appears to be small, and can readily be minimized or removed by averaging the ratings from several judges, or by adjusting the ratings to a form midway between the actual frequency curve and the theoretically correct one

A critical examination of the conditions of the experiment showed that the virtual absence of any personal-equation effect could not have been caused by knowledge of the children's IQ's, nor by discussion of the children among the judges. The inter-school personal-equation effect (between judges in different schools, unfamiliar with each other, and rating different samples) is scarcely any greater than that between judges in the same school.

To minimize the "personal equation" from all sources, the following measures are recommended.

1. *The Judges.* For each subject, ratings should be obtained from several judges. These judges should be interested in the ratings they are making, and have the time to do the rating as well as they are able The judges should be similar in their educational background and professional experience; and they should have some familiarity with the statistical concepts involved in frequency curves If possible, the judges should be given specific training in rating, and the ability of each judge to rate specified traits should be known.

2. *The Rating Scale* The rating scale should state as clearly and objectively as possible exactly what the judges are to rate, and on what observational basis their ratings are to rest. The rating scale should contain very full, quite explicit instructions, with warning against the common errors of rating, viz, "halo", over-rating (especially in cases of long acquaintance and friendly interest), the "relative" effect, and the error of central tendency, or other undue deviations from the theoretically correct frequency curve.

3 *The Statistical Technique* The number of cases rated by each judge should be sufficiently large to make the computation of the mean and standard deviation of his ratings practical Raw ratings should be converted into sigma scores; and if possible, the mean, SD, and skewness of each judge's ratings adjusted on the basis of what-

ever independent theoretical or practical information is available[10] The adjusted sigma-ratings of several judges on each subject should then be averaged—a weighted average is, of course, best if the relative abilities of the several judges are known

It is realized that the set of recommendations above is more or less idealistic, and cannot always be fulfilled Complete fulfillment, however, is by no means necessary in order to keep personal-equation effects negligibly small As a matter of fact, in the present experiment most of the suggested precautions were deliberately violated or ignored—with what results the reader is already familiar

We should not care to deny that situations can be imagined in which the personal equation between judges may become of some importance But it is also true that equally practical situations can as readily be imagined in which the error from the personal equation may be expected, on the basis of the data of this paper, to be negligible So long as correlations with criteria are in the neighborhood of (say) 75 or below, other extraneous factors—either in the criteria, or in the instrument of measurement—are of so much greater importance that the personal equation sinks into relative insignificance. This study seems to demonstrate conclusively that if ratings are to be rejected as a legitimate technique in scientific psychology, they cannot at present be rejected on the ground of the "personal equation."

References

1 Conrad, H S The bogey of the personal equation in ratings of intelligence *J Educ Psychol*, 1932, **23**, 147-149

2 ———— The personal equation in ratings II. A systematic evaluation *J. Educ Psychol*, 1932, **23** (in press).

3 ————. The adjustment of frequency distributions *J Educ Psychol*, 1930, **21**, 386-387

4 Kelley, T L Statistical method. New York: Macmillan, 1924 Pp xi+390

5 Kneeland, N That lenient tendency in rating *Person J*, 1928, **7**, 356-366

6 Miller, W S The variation and significance of intelligence quotients obtained from group tests *J. Educ Psychol*, 1924, **15**, 358-366

7 Rhodes, E C On the relationship of the condition of the teeth in children to factors of health and home environment In Vol VII, *Eugenics Laboratory Memoirs* Cambridge Cambridge Univ Press, 1921 Pp 14-21

[10]For example, in the case of ratings of intelligence, such items as the following would be useful education, scholarship, age at leaving school, occupation of father, social status of parents, and the tendency of the frequency curve toward normality

8 SANGREN, P V Comparative validity of primary intelligence tests, *J. Appl Psychol*, 1929, **13**, 392-412

9 THORNDIKE, E L. The measurement of intelligence New York: Teach Coll, 1927 Pp xxvi+616.

10 WEBB, E Character and intelligence. *Brit. J Psychol, Monog. Suppl*, 1915, 1 Pp viii+99

11. YERKES, R M [Ed] Psychological examining in the United States Army. (*Memoirs Nat. Acad. Sci*, Vol. 15) Washington Gov't Printing Office, 1921 Pp vi+890

12 YULE, G U An introduction to the theory of statistics London Griffin, 1924 Pp xv+415

13 ———— Notice of Heron's memoir on *The influence of defective physique and unfavorable home environment on the intelligence of school children J Roy Statis Soc*, 1910, **73**, 547-551

Institute of Child Welfare
University of California
Berkeley, California

L'ÉQUATION PERSONNELLE DANS LES ESTIMATIONS I UNE DÉTERMINATION EXPÉRIMENTALE

(Résume)

Il y a longtemps que l'on critique estimations à cause de "l'équation personnelle" entre les personnes qui les fait, bien que l'on n'ait défini que rarement "l'équation personnelle" elle-même, ni n'ait jamais précisément déterminé son influence Dans cette etude, "l'équation personnelle" signifie les fausses différences entre les moyens ou les déviations ordinaires des estimations de deux juges ou plus, l'*influence* de l'équation personnelle est définie comme l'influence de ces fausses différences sur la corrélation entre les estimations et un critérium adéquat On a fait cette étude dans le but de déterminer l'effet de l'équation personnelle entre les estimations des juges, dans des conditions faites pour élever cette influence au maximum. Dans ce but, on a employé des estimations de l'intelligence par sept maîtresses d'une école maternelle, comme mesure de la valeur des estimations, on a employé une moyenne de 5 à 9 QI dans des tests d'intelligence étalonnés. L'analyse des résultats montre que, malgré les conditions statistiques et expérimentales défavorables imposées exprès, l'influence moyenne de l'équation personnelle, en termes de corrélation, n'est que de 0,017 Le plus petit effet de l'équation personnelle est de 0,000, le plus grand, de 0,064 On ne peut à peine s'attendre à de plus grands effets de l'équation personnelle que dans des circonstances très spéciales On tire donc la conclusion que, s'il faut refuser les estimations comme technique légitime de la psychologie scientifique, on ne peut pas les refuser à présent à cause de l'équation personnelle

CONRAD

DIE EINWIRKUNG DER PERSÖNLICHKEITSUNTERSCHIEDE AUF RANGORDNUNGEN: EINE EXPERIMENTELLE BESTIMMUNG

(Referat)

Rangordnungen (ratings) sind schon lange auf Basis der Unterschiede der Persönlichkeit (personal equation) der verschiedenen Ordnern (raters) kritiziert worden, obwohl die Einwirkung der Persönlichkeit selbst nur selten definiert und nie exakt bestimmt worden ist. In der gegenwärtigen Schrift meint man mit "Persönlichkeitsunterschieden" ("personal equation") fremde (extraneous) Unterschiede zwischen den Durchschnitten oder Normalabweichungen (standard deviation) der Ordnungen zweier oder mehrerer Ordnern (judges), die *Einwirkung* dieser fremden Unterschiede wird definiert als eine Einwirkung auf die Korrelation zwischen den Ordnungen und einem angemessenen Kriterium (criterion). In der gegenwärtigen Schrift beabsichtet man, die Einwirkung der Persönlichkeitsunterschiede unter Bedingungen zu bestimmen, die mit der Absicht festgesetzt worden sind, diese Einwirkung bis zu ihrem Hohepunkt zu steigern. Zu diesem Zweck verwendete man von sieben Vorschullehrern (nursery-school teachers) berechnete Intelligenzberechnungen (intelligence ratings), als Malssstab der Validität verwendete man den Durchschnitt der an 5 bis 9 anerkannten (standard) Intelligenzprüfungen erzielten Zahlen. Eine Analyse der Befunde zeigte, dass die durchschnittliche Einwirkung der Persönlichkeitsunterschiede, in einer Korrelation ausgedrückt, trotz der absichtlich aufgelegten ungünstigen statistischen und experimentellen Bedingungen nur .017 beträgt. Die kleinste Persönlichkeitsunterschiedswirkung beträgt .000, die grösste .064. Grössere Einwirkungen der Persönlichkeitsunterschiede sind ausser unter ganz besonderen Umständen kaum zu erwarten. Man schliesst hieraus, dass wenn Rangordnung als ligitimes verfahren in der wissenschaftlichen Psychologie zu verwerfen sind, so kann diese Verwerfung gegenwärtig nicht auf Grund der Persönlichkeitsunterschiede stattfinden.

CONRAD

A STUDY OF CERTAIN SELECTIVE FACTORS INFLUENCING PREDICTION OF THE MENTAL STATUS OF ADOPTED CHILDREN

OR

ADOPTED CHILDREN IN NATURE-NURTURE RESEARCH*

From the Institute of Child Welfare and Sociology Department, University of Minnesota

ALICE M. LEAHY

I. THE SIGNIFICANCE OF SELECTION IN NATURE-NURTURE STUDIES

Ever since the time of Galton, interest has grown increasingly in the problem of nature versus nurture The relative potency of the two forces in the behavior of human beings remains unanswered There are few, if any, advocates of the dominance of nature who do not recognize the importance of nurture, and, conversely, there are few, if any, advocates of the dominance of nurture who do not recognize the importance of nature. However, because of the significance of the differential effects of heredity and environment in our educational system, efforts should be continued to secure a more exact conception of the rôle of each. In recent studies of the problem, foster children have been used as subjects. But here the question of a fortuitous mental resemblance between foster parents and children arises. If it obtains, clean-cut measurements of heredity or environment are impossible

Do factors of selection operate in the placement of dependent children in foster homes? How consistently are children of superior mentality placed in homes of superior intellectual level, how consistently are children of medium mentality placed in homes of medium intellectual level, etc ? The problem is involved On the one hand, it presupposes an analysis of children according to some standard, and, on the other, an analysis and classification of the homes in which they are adopted.

Modern Child Placement. As child placement work is organized

*Accepted for publication by John E Anderson of the Editorial Board and received in the Editorial Office, December 22, 1931

today, adoptive parents are rarely the initial selectors. Only certain children are presented to them from whom they may choose. Although some children are placed independently of the supervision of any agency, such children are not likely to appear in a nature-nurture research, since investigators are usually dependent on child-placing agencies if they are to secure an adequate population Indiscriminate placement is practically extinct It is indeed a question whether the frankly commercial "maternity-child-placing hospitals" which are known to exist in certain sections of the United States proceed in such a fashion "Anything, so that it is a baby" would not satisfy their patrons

"Fit the home to the child" has been the exhortation in the literature for the past two decades. As early as 1910, Hart (14, p 236), addressing a group of social investigators, spoke to this effect "You must bear in mind that there are first class, second class, and third class children, and there are first class, second class and third class homes."

Slingerland (19, pp 120-123), writing in 1918, emphasized the importance of selective placement "It is important that the right child be selected to fill an approved application Without the expert study of the home no one knows the type of child likely to win a permanent place there." At this time, also, the part that the agencies should play in determining placement was expressly stated by him "The wise social worker will be able to select the group of children shown, and in many cases direct attention to and influence the choice best fitted to enter that home"

In 1930, the responsibility of the placement agency is more strongly stated "To subject a child to a lifetime of failure to meet requirements of his adoptive family and community is a refinement of persecution. A misplaced child, either over-placed or under-placed, often becomes the center of a tragedy" (17)

What does "fitting the home to the child" connote for selection? Obviously, drawing the home that is best suited for the child, i.e., one in which he will be loved and well cared for, one that is able to surround him with the most favorable conditions in which to realize his potentialities. As much as we might philosophize on the merit that accrues to the individual who meets difficulties and overcomes them, the practical wisdom of men would not permit the placement of children into situations which give anything but the maximum promise for happiness to both the child and the adoptive

parents. On a purely rational basis, the gamble is least in those cases in which the disparity in intellectual equipment between adoptive parents and child is small.

From the systematic procedures known to exist in placement work and from the standards that have been reiterated by child welfare leaders, it seems fair to assume that an attempt is made to select children for adoption on the basis of their mental resemblance to their prospective parents. How nearly it is accomplished awaits demonstration

Studies in the Field Selective placement of children in adoptive homes has not been previously investigated. However, the studies of Burks (3) and Freeman (9) of the influence of environment on test intelligence of foster children should be discussed, since both considered the problem in their experimental populations and both advanced arguments of its non-existence. On the probability of its absence their respective conclusions stand or fall

Burks' and Freeman's studies were nature-nurture investigations concerned primarily in measuring the influence of environment on test intelligence They were undertaken at the same time, but pursued independently—one at Stanford University and the other at Chicago.

As indicated by the title of Burks' study, resemblances between foster parent and foster child were compared with those found in a control group of true parent and true child For the latter group the coefficients or correlation were consistent with those established for filial resemblance, i e, about .50 In the case of the foster group the degree of resemblance between parent and child was strikingly low, ranging from less than 10 to .25

From this evidence the author infers that the intelligence of foster children is only slightly influenced by environment The logical force of this inference hinges on the freedom of her foster group from the probability of selective mental resemblance in placement If mental resemblance between foster parents and children obtains, then the relationship exhibited may be due to it rather than to any effect of environment.

Burks circumscribed the composition of her foster group by limiting it to children who had been adopted previously to the age of 12 months, thus minimizing the possibility of precise judgments of mental ability on the basis of overt behavior Further, she restricted the population to the white race and excluded southern European and

Jewish children The imposition of these limitations left only the possibility of selective mental resemblance on the basis of cultural status

One of the author's main arguments for the absence of selective mental resemblance on the basis of cultural status rests on zero correlations between certain indices of foster home ratings and occupational status of 42% of the true fathers Zero correlations are not surprising when one considers the homogeneity in the two populations with respect to occupation, and, further, the relationship of occupation to cultural home rating. In a 10-category occupational classification, the foster fathers concentrate in the 5 highest classes while the preponderance of true fathers is in the 6 lowest groups. The application of Barr ratings on which the correlation was computed would not greatly change these distributions. As will be shown later, male occupation is in part a function of age (mean of true fathers was 29, estimated age of foster fathers at adoption of child was 38), and cultural home rating is definitely allied with occupation. A comparison of occupational status of true paternal grandfathers and foster fathers, or a comparison of the educational attainment of true fathers and foster fathers, if such data were available, would reveal more clearly the presence or absence of selective placement on the basis of cultural status.

Due to the elimination of undesirable and defective children, the possibility of an increased level of intelligence in her foster group was considered by Burks The probability of reduced variability in IQ and, hence, lower correlations for IQ and culture index of the foster home as a consequence, was not considered

Freeman's method was essentially the same as Burks' His foster population differed considerably, however It was practically double in size (401), age at placement ranged from 6 months to 17 years, 6 months, with the mean at 4 years, 2 months, 8 47% were negro children, and four years' residence in the foster home in which the child was located permitted admission into the experimental group.

In contrast to the low coefficients of correlation between test intelligence and foster home rating found by Burks, Freeman secured coefficients that ranged from 32 to 52 when certain sub-classifications were used, and 48 ± 03 for his entire population From this evidence he propounds the influence of environment on mental ability. An influence, commensurate with that established for true parent and child in which both heredity and environment are operative, follows

from the magnitude of his coefficients, providing the data are free from the effect of selective placement.

The possibility of superior foster parents selecting initially bright children was considered a small factor in the relationship, by the author. Arguments were advanced in support of a negligible selective influence since: (1) in 82% of the cases no mental test had been given before placement; (2) adequate family histories of children were not available by which their mentalities might have been estimated; (3) the average age (4 years, 2 months) at which these children were placed in foster homes precluded dependable estimates of mental ability from observable behavior; (4) when negro children were omitted, the coefficient of correlation between IQ and foster home rating remained practically the same; (5) when 156 children, for whom it would have been least possible to estimate their intelligence since they had been placed under the age of 2 years, were considered independently, a coefficient of $.52 \pm .04$ was secured between test intelligence and foster home rating; (6) when 59 children for whom no histories were available were considered separately from the entire group, a correlation of $.51 \pm .06$ between test intelligence and foster home rating was obtained.

In considering Freeman's objection to the possibilities of selective placement, one would promptly agree upon the limitation of estimates of mental ability in which no objective test data enter. At age 4 years, however, the possibility of estimates of mentality from observable behavior is not remote, especially with the children who have been under observation for any length of time. According to the author's tables, the legitimate children, who constituted 65% of the total foster group, were known to the placement agency and cared for in its receiving home or in a temporary boarding home before placement in permanent foster homes for a period of 11 months, on the average. The illegitimate children, who constituted 35% of the group, were known and under care for a period of 3 months, on the average. The mean age at permanent placement for the legitimate adopted children was 5 years, 8 months; for the illegitimate children, 1 year, 7 months.

Whether the family histories in the Freeman population were sufficiently complete to forecast the potentialities of the children can be seriously doubted. Nevertheless, the amount of information reported caused the author to conclude that "the foster children came from inferior homes and had a heredity which was decidedly poor."

It is possible that these same data supplemented by other unrecorded information influenced the placement agency in its choice of foster homes. In so far as this occurred, the reported coefficient of correlation between test intelligence of children and foster home rating is weighted by selective placement.

If all sources of selective placement were removed in the Freeman population, then the correlation $.48 \pm .03$, found for test intelligence and foster home rating when negro children were omitted, would stand as substantial evidence in support of the influence of environment on the mental development of children. The influence of family history and observable behavior, in a portion of the group sufficiently large to affect the coefficient for the entire population, however, is not improbable.

It is apparent from our analysis of the foregoing studies, that a knowledge of selective placement is essential to the interpretation of nature-nurture findings in which foster children are the subjects. Conclusions cannot be made for either nature or nurture, unless it can be demonstrated that the experimental population is free from the effects of selective mental resemblance.

II. Scope of Investigation and Description of Data

Purpose of Present Study. The present study seeks to discover from the social case records of unmarried mothers (1) the differences in family background between illegitimate children who are adopted and those who are retained by their own mothers; (2) the differences in family background between illegitimate children placed at a very early age and those placed in adoptive homes[1] at a later age; (3) the degree of relationship that exists between certain cultural indices, namely, education and occupational status in the family background of adopted illegitimate children and similar indices in adoptive homes. In general, all available aspects of family history which would seem likely to carry implications as to the mental equipment of adopted illegitimate children will be presented.

Number and Source. The data are drawn from the records of 11,742 unmarried mothers, known to the child-caring agencies of

[1]Adoptive home is one kind of foster home, foster home is a generic term including adoptive, free, and boarding homes. Adoptive home implies that the foster parents stand in a relationship to the child similar to that of true parents. They are the legal guardians of the child, and the home is presumably permanent.

Minnesota during the period 1918-1928, and on file in the Children's Bureau of the State Board of Control[2] Every living birth that survived the age of 2 years, and for whom the records revealed reasonably certain information as to placement, is included in the analysis. The number for whom such information was available is 9973, or 86.8% of the cases Of this population, 4213 children were retained by their own mothers and 2875 were placed in adoptive homes The difference, as far as permanent placement is concerned, appears to lie between these two groups.

The permanency that can be ascribed to any placement type in which illegitimate children are found is dependent on the age of the children and the recency of information concerning them. No child was less than 2 years old when his record was read and transcribed, 95% of the children were 3 years old or over. The mean period of supervision was 2 years, 5 months, computed from the date of opening of the record to date of last entry, in a sample of 1200 cases

It is very unlikely that any large number of children classified in adoptive homes will be transferred to other types of homes with the passing of time Of the children classified as retained by their mothers, 546 had spent varying amounts of time in boarding homes and institutions These children were returned to their mothers at the median age of 21.37 months Eighty-seven per cent of the retained children have been with their mothers or maternal grandparents since birth. The age at which the adopted children were taken into their permanent homes is shown in Table 1 and Figure 1.

It is evident that, irrespective of year of birth, age at placement is highly skewed The median for the major portion of our population (children born since 1918)[3] is 6.89 months, 25% of this group were

[2]The data are part of an intensive nature-nurture study which is being carried on by the *Institute of Child Welfare, University of Minnesota*, in cooperation with the *Children's Bureau of the State Board of Control*

[3]The children born (1900-1917) are classified separately in order to differentiate them from children born subsequently to the establishment of the Minnesota Children's Bureau in 1918 The children born since 1917 are unquestionably more representative of adopted illegitimate children in general The children born (1900-1917) have become known to the Children's Bureau in the process of legal adoption They are a sample of illegitimate children born in a period when independent placements were easily made. Very little is on record concerning their true parents, therefore they will not appear in our subsequent analysis. Estimates of the completeness of our data for true parents will be made on the population born in the period 1918-1928

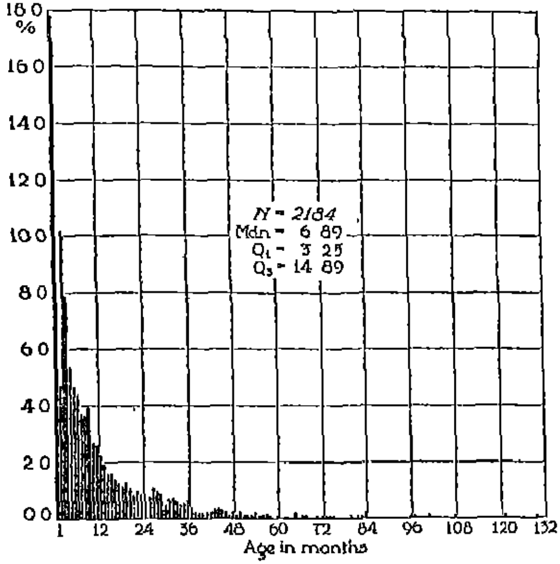


FIGURE 1*A*
BAR DIAGRAMS ILLUSTRATING AGE AT PLACEMENT OF ADOPTED CHILDREN (BORN 1918-1927) TO NEAREST MONTH

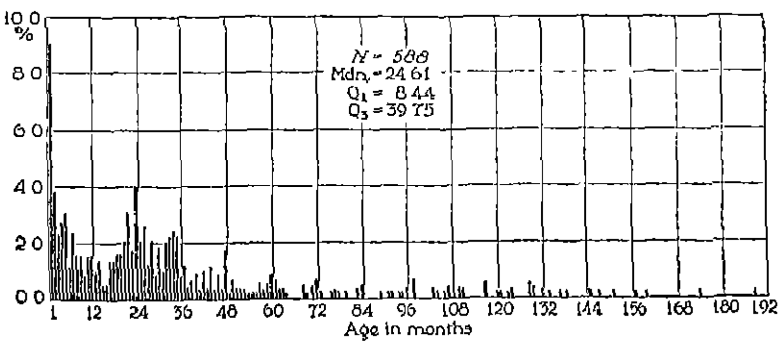


FIGURE 1*B*
BAR DIAGRAMS ILLUSTRATING AGE AT PLACEMENT OF ADOPTED CHILDREN (BORN 1900-1917) TO NEAREST MONTH

TABLE 1

AGE AT PLACEMENT OF ADOPTED CHILDREN (TO NEAREST MONTH)

Birth year	Number reporting age	Percentage of total so placed	Age in months Median	Q_1	Q_3
1918-1927	2184	98.2	6.89	3.25	14.89
1900-1917	588	90.3	24.61	8.44	39.75
1900-1927	2772	96.4	8.10	3.14	21.17

placed at 3.25 months or younger, 75% at 14.89 or younger. That judgments of mental promise, from either test evidence, or observable behavior, could have influenced the selection of homes for the majority of these children is very doubtful. If judgments of mentality were made, they must have been made on other evidence.

Representativeness. To what extent the population that enters into the analysis is unrepresentative of illegitimate children in general cannot be exactly determined. The disappearance of any large number of unmarried mothers for whom the placement of their offspring is unknown would seriously affect our analysis of the relationship of parental background to placement. In view of the relatively small number for whom placement is unknown, and since it is reasonable to expect that this number would be distributed at random in all types of placement (retained by mother, boarding home, institution, and adoption), the likelihood of distorted means is improbable.

An analysis of Minnesota's legal provisions for the registration and supervision of illegitimate children enacted in 1917 (13) leads one to the conclusion that the chances of an unselected population are very great. By these statutes the Board of Control is empowered to demand notification of the birth of all illegitimate children, to license and supervise all maternity hospitals, maternity wards in general hospitals, boarding homes, and institutions for children, and to investigate all petitions to adopt children.

Granting the adequacy of our number, the question of representativeness persists, due to the variability in completeness of the case records. Certain classes of information, as age of parents, residence, marital status, birth date of child, and date of placement, are reported in 60 to 90% of the cases. Education and occupation of mothers are reported in approximately 60% of the records. The same data are less completely reported for the fathers. Data concerning the adoptive parents are 90 to 100% complete. Each class of information, there-

fore, must be considered from the standpoint of its representativeness. The percentage of completeness will be reported on the basis of the total number of children in each placement group. The items *education* and *occupation* will be analyzed on the basis of other known facts, in an attempt to ascertain the existence of possible errors in sampling.

Errors due to the tendency of certain parents to overstate their age, occupation, education, etc., are no doubt cancelled by a similar number who tend to understate these items.

From Table 2 it is apparent that errors due to standards of collecting and recording data would be approximately the same in both populations (adopted and retained), since, judging from birth year of children, neither population is concentrated at any one point in the time span which our data covers. Both populations are distributed with relative equality over the 10-year period. Further, we have no reason to suppose that individuals of a certain social status are portrayed with greater accuracy from population to population. Information concerning the unmarried parent who is most capable of self-maintenance is probably absent in about the same proportion in each population. A comparison of the means of the two groups may hence be valid, even though each of the means entering into the comparison is in error by an unknown amount.

The general tendencies observed in the comparison of these two groups of illegitimate children merit considerable confidence. They are in the main made from a large number of cases appearing over an

TABLE 2

DISTRIBUTION OF ADOPTED AND RETAINED CHILDREN ACCORDING TO YEAR OF BIRTH

Year of birth	Children adopted $N=2249$ Percentage	Children retained $N=4079$ Percentage
1918	10.8	6.5
1919	10.3	5.9
1920	11.3	8.1
1921	12.1	10.0
1922	10.7	11.3
1923	10.2	11.4
1924	9.2	12.2
1925	9.0	11.3
1926	9.5	11.6
1927	7.0	11.6

extended period of time. The large number obviates the possibility of spurious influences from extreme variations, as well as the danger of concentration at any one point in the distribution, for the variable under consideration. The time span involved reduces the influence of transitory social and economic conditions, such as war, economic boom, or economic depression, which might unduly weight our social criteria to one side or the other.

III. Educational Attainment of Parents of Illegitimate Children

The educational attainment of an individual is conceded to be the product of many factors. "The theory that native intelligence is one of the most important conditioning factors in continuance in school" is shown by the accumulation of data on psychological examining in the United States Army (23, Chap. 10). In an unselected group of 653 men from nine camps, the correlation between examination alpha and reported schooling was .75; for a selected group of native-born men, the correlation was .65. The same data, however, demonstrate that education is not "the chief conditioning factor in scores on examination alpha" by a comparison of 660 officers whose schooling was eighth grade or less with 13,943 recruits who had more than eighth-grade schooling. The median alpha score for officers was 107.3, for recruits 97.4. Hence, the Army evidence would advise against unqualified inferences as to the intelligence of an individual from a report of his school attainment. Studies of American school children have shown that individual graduates vary widely from the group averages in test intelligence—attainment being, indeed, the result of factors other than intelligence. However, these same studies have shown the successive elimination of the less well endowed as the educational ladder is ascended. The general expectancy, in considering a group of individuals, is for ability to correlate with length of schooling. And, in comparing groups of indivdiuals, the greater their homogeneity with respect to sex, age, race, community facilities for schooling, and occupation of parents, the more significant are the implications of differences in school attainment for innate ability. An examination of our comparative populations (parents of adopted and parents of retained children) on the foregoing items would obviously yield nothing on subjective factors of personality, educational ideals, or community standards—all of which are unquestionably significant in the school attainment of individuals. However, it is reasonable to

expect that neither population would be weighted more favorably than the other with respect to these subjective characteristics.

Homogeneity of Comparative Population. The *sexes* are considered separately in all comparisons.

There is no difference in *age* for mothers. Age computed to the nearest whole year at the birth date of the child showed the following:

M 21 46, *S D* 4.69, for 1879 mothers whose children were adopted,

M 21 43, *S D* 5 01, for 3958 mothers who retained their children.

A slight difference in mean age was found for fathers

M 25.78, *S.D.* 7.28, for 1300 fathers whose children were adopted,

M 26.46, *S.D* 7 72, for 2929 fathers whose children were retained by mothers.

Notwithstanding the fact that the difference for fathers approaches a value which may be regarded as statistically reliable $\frac{D}{\sigma_{diff}} = 2.79$, the actual discrepancy is not sufficient to raise or depress any educational average that might be explained on the basis of varying educational opportunities with time Certainly, nothing less then a decade would be considered productive of variation in educational attainment due to differences in school opportunities

The percentage frequency distribution of *race* shows that the mothers of children who are adopted are 98 6% white race (N=2035), the fathers, 99 0% (N=1916), the mothers of retained children are 96 1% white race (N=4078), the fathers of retained children are 97.6% white race (N=3886) From these figures it is apparent that the adopted population exceeds the retained in white race. This excess (2%) is so slight, however, that it would not affect an educational average based on a large population A second consideration which makes race a negligible influence in our comparative populations is the absence of information concerning the educational attainment of Indian and negro unmarried parents Ordinarily they were classified under "no information" Therefore, when we contrast placement groups, we are dealing almost exclusively with parents of white race

The major *nationality* groups (German, Norwegian, Swedish) occur in almost identical proportions in the two populations (adopted and retained). In no instance is the disparity greater than 4%—the

average disparity is less than 1%. The group as a whole might be roughly characterized as North European Hence, differences in educational attainment cannot be assigned to differential group traditions

Information on *residence*[4] was available for about 75% of the entire unmarried parent population A statistically reliable difference (7 2%) in favor of city residence in contrast to rural residence is present for the parents of adopted children as compared with parents of retained children There is no difference in town residence, i e., the parents of children who are adopted and the parents whose children are retained by their mothers reside in approximately the same proportion in towns The effect of the difference favoring city residence should be expected to result in higher school attainment for the parents of adopted children. An equally pertinent consideration, however, is the extent to which this discrepancy in residence may be ascribed to innate ability differences of earlier generations The mass of evidence from direct studies (1, 2, 7, 21) leads one to the conclusion that there has been a continuous migration of the more superior intellects to urban communities, since a series of well-controlled investigations show that the intelligence of urban school children is higher than the rural average and that the parents of children of exceptional mental ability are born in greater proportions in urban communities On these bases differences in school attainment between urban and rural population have strong heredity implication. However, with the advent of consolidated schools and transportation at public expense difference in residence is less significant in school attainment. Then, too, we should not overlook the fact that the majority of these parents would have come under the compulsory school law of Minnesota passed in 1899. Compulsory eighth-grade education dates from about the same time or earlier in the neighboring states The residence of 93% of our population is reported as Minnesota and bordering states

In considering the *occupational status* of the parents of our unmarried parents (grandfathers of illegitimate children) we are limited to information concerning 20% of the entire maternal population

[4]Residence was classified according to population, city, town, rural City was defined as a community of 10,000 or over, town, 1000 to 10,000; rural, all under 1000. *The Literary Digest Atlas of the World and Gazetteer 1927*, in which the figures are based on the census of 1920, and estimates supplied by the Bureau of the Census of the population as of July 1, 1925, was used.

and 10% of the paternal population. These data show a statistically reliable difference for high occupational attainment (method of occupational classification presented later) in the case of the maternal grandfathers of children who are adopted, in contrast to the maternal grandfathers of children who are retained The direction of the difference is the same for the paternal population, but is not reliable On the environmental hypothesis, higher educational attainment would be expected for the mothers of adopted children. It should be borne in mind, however, that high occupational attainment of fathers and high educational attainment of children may be due to a factor of heredity common to both parent and child

To summarize, the mothers of adopted and retained children are similar in age, as are also the fathers, the parents of both groups of children resemble each other in race and nationality, the parents of adopted children reside in somewhat larger proportions in cities, in contrast to parents of retained children, the occupational status of the maternal grandfathers of adopted children is higher than that of the maternal grandfathers of retained children

Education of Mothers of Illegitimate Children Education is reported for 1218, or 53 3%, of the mothers whose children were placed in adoptive homes, and for 2848, or 67 6%, of the mothers who retained their children. A comparison of the educational attainment of the two groups of mothers is shown in Table 3.

These data show that a greater proportion of mothers whose children are placed for adoption attain higher levels of education than

TABLE 3

PLACEMENT OF CHILDREN ACCORDING TO EDUCATION OF TRUE MOTHERS

Education of mother	Children adopted No of cases	%	σ_p	Children retained No of cases	%	σ_p	Critical ratio $\frac{D}{\sigma_{diff}}$	Chances in 1000 of a true diff. > zero
Less than high school	694	56 9	1 4	2001	70 3	0 9	7 99	1000
High school not completed	328	26 9	1 3	619	21 7	0 8	3 47	1000
Completed high school or more	196	16 1	1 1	228	8 0	0 5	6 70	1000

do mothers who retain their children. It would appear that the chances of a difference in the same direction as that observed are practically 1000 in 1000 for similarly chosen populations.

Education of Fathers of Illegitimate Children. Education is reported for 178, or 7 8%, of the fathers whose children were adopted and for 352, or 8.5%, of the fathers whose children were retained by their mothers Table 4 presents a comparison of the educational attainment of the two groups of fathers.

Here, as in the case of the mothers, a greater proportion of fathers whose children are adopted attain higher levels of education than do fathers of children retained by their mothers. The probability of a true difference in the same direction as that found is 998 in 1000. However, the number of data concerning the fathers is small when compared with the total population It is obviously less dependable than our data for mothers A correlation of 486± 04 (N=912) between education of fathers and that of mothers (irrespective of placement of children) would cause one to expect the general difference observed for higher educational attainment of mothers of adopted children as compared with mothers of retained children to be maintained for fathers also The sampling errors of our educational data will be discussed in the next section

Representativeness of Data The confidence that can be placed in the foregoing differences in education of mothers and fathers of illegitimate children is obviously dependent on the representative character of our samples. Are the parents whose education is known

TABLE 4
PLACEMENT OF CHILDREN ACCORDING TO EDUCATAON OF TRUE FATHERS

Education of father	Children adopted No of cases	%	σ_p	Children retained No of cases	%	σ_p	Critical ratio $\frac{D}{\sigma_{diff}}$	Chances in 1000 of a true diff > zero
Less than high school	58	32 5	3 5	154	43 8	2 6	2 40	992
High school not completed	43	24.1	3 2	92	26 1	2 4	48	685
Completed high school or more	77	43.3	3 7	106	30 1	2 3	2 97	998

typical of the mothers and fathers of adopted and retained children? Only a knowledge of the complete population with respect to education would answer the question with certainty A critical analysis of the data with respect to the influence of other social criteria, however, would provide a partial test of representativeness. It is conceivable, for example, that constant errors due to residence may be operating. If the group who reports education resides in preponderantly larger numbers in the city, in contrast to the group for whom education is not reported, and if city residents tend to attain higher levels of schooling, then we may conclude that our data are positively weighted by the factor of residence. If there is no difference in residence, it would appear that the educational attainment of the group whose schooling is unknown would not be essentially different from that presented by our findings, in so far as residence is concerned In considering the two populations (adopted and retained) the method of checking against other social factors sharpens the implications of our evidence Sufficient data are available to make possible a comparison on the basis of age and residence between individuals whose education is given and those for whom education is not reported Although we are able to check only a limited number of the variables related to school attainment, and only a portion of our entire population, the portion for whom age and residence are known is exceedingly large

Comparing mothers whose education is known and mothers whose education is unknown, it was found that the mothers in both sets of data (adopted and retained) whose education is unknown are older—the chances of a reversal in mean ages in each instance are less than 1 in 1000 A low negative correlation ($-.229 \pm .01$, $N=7135$) between education and age indicates that our educational data are positively weighted for the mothers in both adopted and retained populations. When residence was considered no difference was found between the mothers whose education was reported and the mothers whose education was not reported Therefore, notwithstanding the probability of lower educational averages if complete information were available, there is no evidence of a reversal of the difference in education between mothers of adopted and retained children

A similar analysis of fathers whose education is known and fathers whose education is unknown shows no differences for age or residence in the adopted population In the case of fathers of retained children, the group for whom education is unknown is re-

liably older than the fathers whose education is known. Thus we might expect a lower average education for the entire retained population since age and education of fathers correlate $-.321 \pm .03$ ($N=921$) A lower average education is also suggested by the direction of the difference for rural residence Hence the true difference between the education of fathers of adopted and retained children would likely be greater than that presented in our data.

From the available evidence we can conclude that a larger proportion of mothers and fathers of children who are relinquished for adoption attain higher levels of education than parents whose children are retained by their mothers Although one would be inclined to hold this conclusion as tentative in the case of fathers, it is clearly decisive for mothers, where the number of the data eliminate the likelihood of the difference in education arising from chance

The theoretical significance of the above observation is highly important in nature-nurture investigations involving adopted children. The usual procedure is to assign the differences in actual test intelligence of adopted children and expected intelligence made from estimates of parental intelligence to the influence of environment It is apparent that the amount assigned to environment will be determined by the general figures for intelligence of true parents that prevail. To date, studies indicate that unmarried mothers are below the average in mentality and in cultural background But, the fact should not be lost sight of that these studies (22) have been limited to small populations of selected cases known to social welfare agencies and drawn from areas where the unmarried mother who relinquishes her child for adoption is practically unknown From our data it appears that true mothers of adopted illegitimate children are at the upper end of the distribution of unmarried mothers in cultural background as judged by education, therefore, the probability that they would vary far from the average mentality of the general population is unlikely Regardless of the need of evidence on a variety of attributes before a general characterization as to level of ability can be made, one is apparently not justified in assuming an arbitrary level of inferiority for these mothers or their relinquished children. In view of the homogeneity observed in the two populations (adopted and retained), the superiority of mothers of adopted children in school attainment suggests the probability of relative superiority in mentality for this group as a whole

IV. Occupational Status of Parents of Illegitimate Children

Among the measures of economic status, paternal occupation has been demonstrated to be related to a variety of mental traits in children. Haggerty and Nash (12), in a study of 8121 New York State school children in Grades 3-12 inclusive, found that success in intelligence is directly related to the occupation of the fathers. Similar intellectual differences between occupational classes were found by Goodenough (10, pp. 45-57) in children two to four years of age. Gifted children, as defined by Terman (21, pp 61-72), were found in preponderantly larger numbers in the homes of parents of the higher occupational levels. The relation between occupation of parent and intelligence of children was shown to be positively correlated by Dexter (6) in a study of 2782 school children, Grades 1-8 inclusive When the social origin of orphan children was rated on a 5-category occupational scale in Jones's (15) study of 831 orphans, the children from the two highest occupational classes secured significantly higher intelligence test scores Studying a large population of seniors in the high schools of Massachusetts, Colvin and MacPhail (4) found test intelligence to be positively associated with paternal occupational rank Investigations of the composition of our elementary- and high-school populations have shown that the children in the higher grades are recruited in significantly larger proportions from the upper occupational levels This was clearly demonstrated by Counts (5) in his study of over 170,000 high-school students in four urban communities Sears *et al* (18) also showed this relationship in their survey of the schools of Boise, Idaho As the educational ladder is ascended, the children from the lower occupational levels are eliminated Investigations of courses of studies pursued and success in school work have shown a similar tendency for the children of higher occupational levels to choose the courses that make the greater demand on abstract thinking and, in general, to exhibit more proficiency in the acquisition of knowledge Therefore, it would appear that in so far as we may regard test intelligence, grade in school obtained, and success in school studies, functions of mental ability, a comparative analysis of paternal occupation would tend to differentiate groups of individuals on the basis of mentality While occupation is only a crude measure, and while the tendencies previously recorded show many individual exceptions, they will nevertheless be found to hold, on the average, for large groups Further,

it should be remembered that from data of this kind no inferences can be made as to whether the cause is a matter of heredity or environment To what extent the superior cultural opportunities afforded children of the upper occupational classes tend to increase their score on intelligence tests and to stimulate greater attainment and proficiency in school work, or to what extent the ability of the parents which enables them to attain higher occupations is inherited and may be the cause of the child's performance is unknown

Occupational Classifications. The occupations of fathers and grandfathers are classified on an occupational scale, consisting of a hierarchy of seven occupational groups, constructed from the occupations listed for adult males for Minnesota in the report of the 1920 census The hierarchical arrangement assumes that each successive group (VII to I) makes greater demands on mental ability The occupations are not ranked within the groups, they are simply listed in the group in which, in the opinion of the judges,[5] there is a certain amount of homogeneity from the standpoint of the demand made on mental ability The groups range from those occupations demanding a high degree of training and ability to those exacting no training and very little ability. For example, Group I includes such occupations as chemical engineering, Group VII, day labor In an analysis of investigations made of parental occupation and test intelligence of children, Sorokin (20, p 240) found that in 11 out of 14 studies children of farmers ranked between the children of skilled and semi-skilled laborers. Hence, in the construction of the occupational scale for this study the farmer was classified between the skilled and semi-skilled groups. The positions of the remaining groups correspond fairly well with positions assigned them in other empirical scales. A general description of the occupational groups (11, p 237) follows

Group I	Professional
Group II	Semi-professional and managerial
Group III	Skilled trades, clerical, and retail business
Group IV	Farmers
Group V	Semi-skilled occupations, minor clerical, and minor business
Group VI	Slightly skilled trades, occupations requiring little training or ability
Group VII	Day laborers of all classes

[5]F L Goodenough, Institute of Child Welfare, University of Minnesota. E A. Rundquist, Child Guidance, Minneapolis, Minnesota. A M Leahy, Institute of Child Welfare and Department of Sociology, University of Minnesota

The occupations of mothers are classified on a scale similarly constructed from the distribution of the employed single females in Minnesota in the census report of 1920. Since the occupation of farmer appeared for less than 4% of the entire employed single female population, no separate group was made for them, they are classified with the semi-skilled As a consequence the scale for women consists of six groups These groups are similar to Groups I, II, III, V, VI, and VII of the classification for men.

Scales made from a more exact knowledge of the relationship of occupational attainment to mental ability awaits future investigation.

Occupation of Mothers Occupation during pregnancy is reported for 1308, or 57 2%, of the mothers whose children are adopted and for 2810, or 66 7%, of the mothers who retained their children. Classifying these data on the Minnesota Occupational Scale, Table VI, a consistent tendency for the mothers of adopted children to exceed the mothers of retained children in the higher occupational groups is apparent. If the populations are contrasted on the basis of the percentage frequency in Occupational Groups I, II, and III combined, the chances of a difference in the same direction as that observed are about 1000 in 1000 ($\frac{D}{\sigma_{diff}} = 3.50$) for similarly chosen populations. When occupation at conception is considered the same tendency is observed, although the data are less numerous (N for adopted population = 282, retained = 725) The concentration of both groups of mothers in the lower occupational levels is in part due, unquestionably, to their physical condition which bars employment in higher occupations The largest percentage are in domestic employment, Group VII. Considering, also, the fact that previously married mothers are included with single mothers of illegitimate children, we should expect the distribution to be weighted in the lower occupational groups, since, in general, it has been observed that married women exceed single women in the lower occupational levels Whatever disadvantage is imposed on our populations by classifying them on a scale constructed for single women, however, would be greatest for the mothers of adopted children since their degree of singleness is somewhat less than that of mothers who retain their children, namely, 86% ($N = 1680$) as against 90% ($N = 3748$) In view of the latter, the difference favoring higher occupational attainment for mothers of adopted children is all the more notable

TABLE 5

PLACEMENT OF CHILDREN ACCORDING TO OCCUPATION OF PARENTS

(Occupation of mothers as reported during pregnancy)

	Mothers		Fathers	
Occupational Group	Children adopted $N=1308$ Percentage	Children retained $N=2810$ Percentage	Children adopted $N=1323$ Percentage	Children retained $N=3120$ Percentage
I	0 4	0 1	2 1	1 7
II	4 0	2 9	4 5	3 1
III	8 7	5 9	18 1	13 9
IV			9 0	14 2
V	25 8	22 1	35 8	29.2
VI	22 6	23 5	13 9	14 4
VII	38 5	45 5	16 6	23 5

Occupation of Fathers. As shown in Table 5, occupation is reported for 1323, or 57 8%, of the fathers whose children are adopted and for 3120, or 74 1%, of the fathers whose children are retained by their mothers. It is evident that a larger proportion of fathers of adopted children obtain higher occupations than do fathers of children retained by their mothers. The difference between the respective distributions, on the basis of the percentage frequency in Groups I, II, and III combined, is of such magnitude that the chances of a difference greater than zero are roughly 1000 in 1000 $\left(\frac{D}{\sigma_{diff}} = 4.60\right)$ for similarly chosen populations

Representativeness of Occupational Data While there is little doubt that the large number of cases involved in our occupational data is sufficient to protect it against the influence of variable errors, nevertheless, the nature and direction of systematic errors remains a question deserving examination. To make such an examination adequate, a knowledge of a wide variety of facts concerning the health, intelligence, education, age, residence, etc, of our entire population would be desirable We are limited to an acquaintance with two of these factors—age and residence—as they affect a portion of our population In all instances, however, the number that enters into this examination is more than 60% of the entire population

The analyses of the data in both populations (adopted and retained) show that the mothers whose occupations are known are

older than those mothers whose occupations are unknown. Age of mothers whose children are adopted is

M 21.77, $S.D.$ 4.36 ($N = 1348$), occupation reported,
M 19.86, $S.D.$ 4.73 ($N = 422$), occupation not reported

Age of mothers whose children are retained is

M 22.00 $S.D.$ 4.84 ($N = 2715$) occupation reported,
M 21.03, $S.D.$ 5.47 ($N = 556$) occupation not reported.

From the above it is apparent that the difference in age is greater within our comparative populations than between the samples whose occupations are known. Statistically, the difference, within each population, between mothers whose occupations are known and those whose occupations are unknown is sufficiently large that the chances are about 1000 in 1000 of a true difference greater than zero for similarly chosen populations. No inferences can be made as to whether our occupational distributions have been raised or depressed because of age selection. The relationship of age and occupation for women is unknown. Many of our younger mothers have had no occupational history. All that is important in this analysis is the close agreement in age (21.77 and 22.00) of mothers whose occupations are known. Obviously, the difference in occupational status between mothers of adopted children and mothers of retained children cannot be ascribed to age.

When the *residence* of mothers whose occupations are known and those whose occupations are unknown is examined, no difference is found in the adopted population. But in the retained population, the difference for city residence is reliably greater for those mothers whose occupations are known. Hence, if city residence is favorable to higher occupational attainment, we should expect the occupational average of the entire retained population to be lower than that reported in our data. From this it follows that the difference in occupational status between our two populations (adopted and retained) would, in all probability, be greater than that presented in Table 5 if complete information were available.

Systematic errors in our occupational data due to age, but in the same direction for both populations (adopted and retained), appear on an examination of our data for fathers. Age of fathers whose children are adopted is

M 26.50, $S.D.$ 7.50 ($N=994$) occupation reported,
M 24.04, $S.D.$ 6.18 ($N=208$) occupation not reported

Age of fathers whose children are retained is

M 26.83, $S.D.$ 7.55 (N=2428) occupation reported,

M 23.02, $S.D.$ 9.58 (N=402) occupation not reported

It is evident from the above that our occupational data concern two populations that agree closely in age (26.50 and 26.83), and hence the observed difference in occupational status cannot be assigned to difference in age.

In so far as residence would tend to raise or depress an occupational average, there is not any difference within our respective populations between fathers whose occupations are known and fathers whose occupations are unknown. In no instance is the difference greater than one and one-fourth times its standard error. Therefore, neither occupational distribution (adopted and retained) can be said to be unrepresentative of its population because of residence.

From the available evidence it appears that when our data are examined for sampling errors due to age and residence, the relative superiority in occupational attainment observed for true parents of adopted children, in contrast to true parents of children retained by their mothers, is maintained. This observation, it will be recalled, is made from large samples (over 1000) in which chance differences are improbable. Conceding the limitations of occupational status as a discriminating index of socio-economic status, it is significant that in this study it corroborates the evidence on educational attainment of parents as to the selective character of adopted children.

V. Cultural Status of True Parents and Age at Placement of Adopted Children

Since unmarried parents of relatively superior cultural attainment, as judged by education and occupation, tend to relinquish their children for adoption, the next question of interest is the relationship of age of child at relinquishment and degree of cultural attainment of the parents. Although age at relinquishment does not necessarily coincide with age at placement in an adoptive home, placement generally follows soon after relinquishment. Otherwise, institutions would be crowded beyond capacity, and adoptive parents would be denied the opportunity of rearing children from the earliest age at which they are available. Therefore, any relationships based on either age at placement or age at relinquishment would, in general, differentiate the parents of illegitimate children, notwithstanding

the fact that there are individual exceptions in which a child is held in an institution for several years after relinquishment In Minnesota, as in most states, retention by the mother is the prevailing policy, adoption is discouraged, while institution and boarding-home care are regarded as temporary—awaiting permanent placement with either the mother or adoptive parents

A further aspect of age at placement arises from the side of the adopters. Applications for children tend to show that, of the specified attributes desired in a child, age is among the foremost While the limitations set for age are unquestionably related to certain factors, e.g , wealth of adopters, emotional attitudes, training ideals, etc , their implications for mental resemblance have general scientific importance Freeman (9, p 196) found that the younger the child at age of placement, the more superior was the cultural rating of his adopted home, and the higher his intelligence If children who are placed at very early ages are superior in heredity to those entering adoptive homes at older ages, then his data would indicate the influence of selective placement in mental resemblance

In our data, Table 6, it appears that the education of mothers of children placed at 3 months or under (Q_1) is superior to that of mothers whose children are placed at 21 months or over (Q_3). Further, a coefficient of contingency, computed on 2883 cases, of .339 $\pm$.03 between education of mother and occupation of maternal grandfather leads to the inference that the more superior the maternal social status, the earlier the child is placed for adoption. If social status and heredity are positively associated, it follows that foster parents who adopt very young children are in all probability adopting children who are relatively superior in intelligence to other dependent children. The above observation is maintained for the education of fathers also The possibilities for sampling variance are too great, however, to offer these data as evidence (number of cases 53 and 28)

In Table 7 the occupational attainment of parents whose children are placed at 3 months or earlier is contrasted with parents whose children were placed at 21 months or over

A consistent tendency is apparent in these data for the occupation of parents of children placed at 3 months or earlier to be superior to that of parents of children placed at 21 months or over Notwithstanding the fact that the magnitude of the difference is such that the chances of a reversal are within the errors of sampling, the consistency of its direction is striking when the reduced variability in occupation

TABLE 6
AGE OF CHILDREN AT PLACEMENT ACCORDING TO EDUCATION OF MOTHERS

Education of mother	Child placed at 3 mos or younger Number =317		Child placed at 21 mos or older Number =220		Critical ratio $\frac{D}{\sigma_{diff}}$	Chances in 1000 of a true difference >zero
	Percentage	σ_p	Percentage	σ_p		
Less than high school	52.7	2 8	67 7	3 2	3 57	1000
High school not completed	24 9	2 4	21 8	2 8	0 84	800
Completed high school or more	22 3	2 3	10 5	1 9	3 96	1000

TABLE 7
AGE OF CHILDREN AT PLACEMENT ACCORDING TO OCCUPATION OF PARENTS

Occupational group	Mother: Child placed at 3 mos or younger $N = 372$ Percentage	Mother: Child placed at 21 mos or older $N = 260$ Percentage	Father: Child placed at 3 mos or younger $N = 345$ Percentage	Father: Child placed at 21 mos or older $N = 265$ Percentage
I	0 0	0 0	2 6	1 5
II	5 1	3.8	6.1	4 9
III	9 1	6.2	20 3	18 1
IV	—	—	9 0	9 1
V	28 2	24 6	35 1	34 7
VI	19 9	26 9	10 4	15 1
VII	37 6	38 5	16 5	16 6

for fathers of our age group is considered and the limitations in occupations for pregnant women are borne in mind If a greater amount of data were available on occupations of mothers at conception, differences according to age of child at placement would undoubtedly be more significant, since occupation previous to pregnancy would tend to differentiate true attainment more clearly Our data on occupation at conception are entirely too few to draw any conclusions. However, the mothers of children placed at the earlier age continue to attain higher occupations Superiority in occupational attainment of forebears obtains for the children placed at 3 months or under when the occupations of grandfathers are considered also, but here, too, the number of data is not large enough to warrant a specific conclusion, although the trend of the evidence is consistent with that found for mothers and fathers

VI Selective Placement of Adopted Children

As stated previously, a study of the resemblance between adoptive parents and children in a variety of traits, based on direct measures taken before placement, would undoubtedly be the ideal procedure for an investigation of the question of selective placement. A random population and reliable measures are implicit in the design Securing an adequate population would be the lesser difficulty Reliable tests, however, of simple physical traits such as eye or hair color, dwindle when the early age at which children enter adoptive homes is recalled. Measures of mental characteristics of children, age 6 months, mount high in their unreliability. The nearest approximation to the ideal is a study of resemblance between true and adoptive parents. And here we must be content with resemblance in social status, for in no section of the country is there available objective evidence of mental characteristics, notwithstanding the fact that data have accumulated in certain research bureaus concerning true parents who present difficulties in social adjustment In no sense may such selected groups be regarded as typical of true parents, for they are ordinarily drawn from those giving the greatest evidence of incompetence

Social status, as judged by educational attainment, occupation, home rating, rent, income, property holdings, etc , of true parents and adoptive parents, is perhaps the best criterion by which we may estimate the probability of selective placement. And, although the youth of true parents precludes evidence on any but the first two listed, the

rank of their parents on the latter would seem adequate for determining whether any relationship exists between the social background of the child and the social status of the home in which he is placed. Specifically, the question may be stated as follows. Are children of parents of high social status placed in adoptive homes of high social status, those of medium status, in medium homes, etc ? If superior social status implies superior mental ability, on the average, then inferences may be made for mental resemblance in placement, although the injection of the hypothesis does not alter the evidence Assuming that the hypothesis is correct, our question becomes Are children whose family background gives the greatest promise of mental superiority placed in adoptive homes of superior social attainment, etc.? If a positive answer is found to the question, then we may expect a correlation between the level of the adoptive home and the intelligence of the adopted child, entirely apart from any direct effect of training or environment The Minnesota data leave much to be desired from the standpoint of size of sample in the case of true fathers, and also in the number of variables on which the adoptive and true parents may be compared.

Before discussing the implications of the above coefficients it will

TABLE 8
COEFFICIENTS OF CONTINGENCY BETWEEN THE EDUCATION OF ADOPTIVE MOTHERS AND OTHER FACTORS

Correlated factor	Number of cases	Coefficient of contingency Uncorrected	Corrected	P E
Education of true mothers	836	371	286	04
Education of true fathers	127	485	340	07
Occupation of true mothers at conception	195	408	326	06
Occupation of true mothers during pregnancy	968	183	129	03
Occupation of true fathers	905	408	418	02
IQ of child (K B, not differentiated as to time of test, before or after placement)	93	463	338	.12
Average		386	306	

TABLE 9

COEFFICIENTS OF CONTINGENCY BETWEEN THE EDUCATION OF ADOPTIVE FATHERS AND OTHER FACTORS

Correlated factor	Number of cases	Coefficient of contingency Uncorrected	Corrected	P.E
Education of true mothers	832	241	225	03
Education of true fathers	124	403	.310	07
Occupation of true mothers at conception	196	367	248	12
Occupation of true mothers during pregnancy	971	175	116	03
Occupation of true fathers	1131	174	108	02
IQ of child (K B, not differentiated as to time of test, before or after placement)	98	498	367	13
Average		310	229	

be well to call to mind that a coefficient of contingency (C) is comparable with the product-moment coefficient of correlation "if for each variable the categories are successive values of a graduated variate, and if the population is large and the number of categories great so that there is not a grouping error, and if the correlation surface is normal" (16, p 266). To the extent that successive levels in school may be regarded as graduated we are dealing with a graded series for education, although not in the same sense as a scale of inches, in which any fractional part may be measured and recorded and the multiple of any value established Although the distance between school grades is regarded as 1, the distance in difficulty is not known The graduated character of the occupational variable is somewhat similar to that of education, but more arbitrarily imposed Each successive group from VII to I presumably makes greater demands on mental ability, the distance between groups is unknown There is no question as to the graded character of IQ and age data. In exactly 50% of the relationships established the population exceeds 800, and hence, for these particular coefficients, it would seem that our data satisfy the second criterion Since the number of categories employed in these computations was never less than 25 (the minimum for valid

analysis) the third prerequisite is met. In general, a normal correlational surface is approximated in all the contingency tables.

Referring to Tables 8 and 9, it will be seen that there is a reliable tendency for *education* of true and adoptive parents to be correlated. The correlations for occupation of true parents and education of adoptive parents, although less consistent in size when adoptive parents are considered separately, indicate a tendency for *education* of adoptive parents to be related to *occupation* of true parents Reduced variability in occupation of true mothers during pregnancy undoubtedly accounts for the low coefficient obtained when this variable is correlated with the education of adoptive parents. The degree to which IQ of adopted children and educational attainment of adoptive parents is associated cannot be learned from our data The figure given here is based on a small number of cases, undifferentiated as to time of test.

In general, it appears that the coefficients of contingency between the education of adoptive mothers and factors in the background of adopted children are consistently higher than those secured for the education of adoptive fathers and similar factors in the background of adopted children. The explanation lies possibly in the fact that child-placing agencies ordinarily see more of the adoptive mother than they do of the adoptive father and therefore, in estimating the "fitness" of a home, base their judgments primarily on evidence concerning the adoptive mother.

The evidence in Table 10 gives low positive relationships for occupation of adoptive fathers and the following· education of true mothers, occupation of true mothers during pregnancy, and occupation of maternal grandfathers. Positive, but less reliable, relationships are suggested for the other variables listed with the exception of occupation of true fathers, where the relationship appears to be zero The explanation for the latter probably lies in the selected character of the two populations, viz., the true fathers are concentrated in the lower occupational groups, while the adoptive fathers are concentrated in the higher occupational groups. The possibility that the concentration of true fathers in the lower occupational levels is in part a function of age, and therefore zero or negative correlations are to be expected, finds support in the application of the Occupational Classification (11, pp 501-512) constructed for the adult employed males in the United States to males who are listed by occupation and age, for the entire United States in the 14th census (8, pp 378-395).

TABLE 10

COEFFICIENTS OF CONTINGENCY BETWEEN THE OCCUPATION OF ADOPTIVE FATHERS AND OTHER FACTORS

Correlated factor	Number of cases	Coefficient of contingency Uncorrected	Corrected	P.E
Education of true mothers	937	222	174	.02
Education of true fathers	138	471	269	10
Occupation of true mothers at conception	213	395	.219	08
Occupation of true mothers during pregnancy	1086	199	130	03
Occupation of true fathers	1046	210	080	03
Occupation of maternal grandfathers	307	430	308	06
Occupation of paternal grandfathers	109	485	289	12
Average		344	209	

Since it is probable that the Minnesota employed males do not differ radically in age from employed males in the entire United States, and since the two Occupational Scales agree in the classification of occupations common to Minnesota and the United States, the occupational distributions according to age shown in Table 11 are undoubtedly very similar to what would be found if evidence on age and occupation were available for the general population in Minnesota.

Table 11 shows a clear tendency for younger men, in general, to be concentrated in the lower occupational groups Therefore, we may expect a comparison of the occupations of men who differ widely in age to yield zero or negative correlations The percentage frequency of the true and adoptive fathers in the lower occupational groups differs greatly, 66 3 and 30 2, respectively The occupation of the fathers of the parents of illegitimate children (paternal grandparents of illegitimate children) would undoubtedly be a more satisfactory social datum from which to secure evidence of selective placement It is ordinarily conceded that the occupational attainment of a young woman's father is a better index to her social status than the occupation which she fills. The hypothesis probably holds

TABLE 11

COMPARISON OF PERCENTAGE FREQUENCY DISTRIBUTION OF TRUE FATHERS, EMPLOYED ADULT MALES IN UNITED STATES, AND ADOPTIVE FATHERS BY OCCUPATION

		Employed adult males in U S Total $N = 28,300$ 421			
Occupational Group	True fathers M age 26 50 $N = 1323$ Percentage	Age 20—21 $N = 4,121,914$ Percentage	Age 25—44 $N = 15$ 603,875 Percentage	Age 15—67 $N = 8,574,632$ Percentage	Adoptive fathers M age 39 28 $N = 2290$ Percentage
I	2 1	1 7	3 0	3 2	9 6
II	4 5	2 4	5 3	5 9	15 3
III	18 1	14 1	16 2	15 3	24 4
IV	9 0	9 8	18.6	27 0	20 6
V	35 8	32 0	29 3	23 4	22 4
VI	13 9	15 1	10.2	9 1	5 2
VII	16 6	24 9	17.4	16 2	2 6
Percentage in Groups V, VI, VII	66 3	72 0	56 9	48 7	30 2

for men under 30, also. The coefficient of contingency .308 ± 06 for occupation of maternal grandfathers and the occupation of adoptive fathers is in general agreement with the average of the relationships found for education of adoptive mothers with factors in the background of the adopted child

An analysis of the foregoing data shows that certain factors in the family history of adopted children, namely, education of true parents and occupation of the maternal grandfather, are related to the cultural status of the adoptive home to a degree beyond that which might be accounted for by chance While the occupation of true parents is not of outstanding importance, it cannot be said to be of no importance. If the traits considered were completely unrelated (independent) $C = 00$ A departure from independence is clearly shown The averages, .306, 229, and 209, are not greatly reduced if calculated from only those coefficients that are at least four times their own probable errors. When restricted in this way, they become .299, 189, and .204 If more data were available, an average coefficient of resemblance of about 30 is highly probable.

Notwithstanding the general exploratory character of this investigation and the unreliability of certain of the measures employed, the trend of the evidence points directly to the conclusion that any direct study of children in adoptive homes cannot ignore the probability of selective placement Unless the experimental population is free from the influence of selective placement, the correlations will

present the complex of environment and selection, which makes them analogous to true parent and true child correlations, where the complex is environment and heredity If an experimental population is older at age of placement than the one considered here, and is drawn from an area comparable to Minnesota in its standards of child-placing, the probability of selective placement is undoubtedly greater than that exhibited in our data

Summary

1 The present investigation was undertaken to secure whatever evidence was available in social case records concerning the selection of illegitimate children for placement in adoptive homes The significance of selective placement in a research population of adopted children chosen for the purpose of measuring the influence of environment on mental ability is discussed

2 Factual data concerning the educational attainment and occupational status of true and adoptive parents are examined in a total population of 2287 illegitimate children placed in adoptive homes and in 4213 illegitimate children retained by their own mothers The data were secured from records on file in the Children's Bureau of the State Board of Control, St Paul, Minnesota These records cover a 10-year period, 1918-1928. Due to the inclusive character of the Minnesota legislation for the registration and supervision of illegitimate children, they are perhaps the most unselected sample of records of illegitimate children ever assembled for analysis

3 A report of educational attainment was available for 1218 true mothers of adopted children and for 2848 true mothers of retained children. A comparison of these data classified on a three-category basis, *less than high school, high school not completed, high school completed or more,* shows a reliable difference in favor of higher educational attainment for mothers of adopted children. A similar tendency is observed when true fathers of adopted children ($N=178$) are compared with true fathers of retained children ($N=352$) On number alone it would appear that there is considerable possibility of sampling errors in the data for fathers An examination of both sets of data (fathers and mothers of adopted children, fathers and mothers of retained children) with respect to age and residence indicates the probability of greater differences than those exhibited The implications of differences in educational attainment for differences in mental ability are discussed and the

fallacy of assuming that true mothers of adopted children are similar in mental equipment to a selected population of unmarried mothers known to social welfare agencies is pointed out.

4 A report of occupation for 1308 mothers and 1323 fathers whose children were adopted and 2810 mothers and 3120 fathers whose children were retained by their mothers shows that a reliably greater proportion of parents of adopted children attain higher occupations than do parents of retained children In no instance could the difference be ascribed to age or residence.

5. The relationship of age of child at placement to education of true mother discloses the fact that mothers of children placed at 3 months or younger attain a reliably higher educational rank than mothers whose children enter adoptive homes at 21 months or older A consistent tendency is apparent also for the parents of children placed at the earlier age to be superior in occupational attainment In the light of these findings and that of Freeman's of an inverse relationship between level of adoptive home and age of child at placement, the probability of a positive correlation for mentality of adopted child and level of adoptive homes is considered.

6. The degree of resemblance found between the cultural level of adoptive homes and the cultural level of true parents of adopted children is expressed by the following coefficients of contingency.

.306 for education of adoptive mother,

.229 for education of adoptive father, and

.209 for occupation of adoptive father, with factors in the family background of adopted children. On the hypothesis of a positive correlation between cultural status and mentality level, our results indicate that there is a resemblance between adoptive parent and child entirely apart from any influence of environment and training

References

1 Baldwin, B T, *et al* Farm children New York Appleton 1920 Pp 337

2 Book, W F Variations of mental ability and its distribution among the school population of an Indiana county *Bull Exten Div, Ind Univ*, 1918, 4, No 4, 100-140

3 Burks, B S The relative influence of nature and nurture upon mental development A comparative study of foster parent–foster child resemblance and true parent–true child resemblance *27th Yrbk, Nat Soc Stud. Educ*, 1928, Pt 1, 219-316.

4. Colvin, S S, & MacPhail, A H Intelligence of seniors in the high schools of Massachusetts *U S Bur Educ., Bull.*, 1924, No 9 Pp 39

5 COUNTS, G S Selective character of American secondary education. *Suppl Educ Monog*, 1922, No 19 Pp 162

6 DEXTER, E S The relation between occupation of parent and intelligence of children *School & Soc*, 1923, 17, 612-614

7 ENGELHART, F Survey of Lake Crystal schools. *Bull Minn Coll Educ*, 1926, Monog 10 Pp 100

8. Fourteenth Census of the United States Vol. 4 Pp 951-955

9 FREEMAN, F N, HOLZINGER, K J, & MITCHELL, B C The influence of environment on the intelligence, school achievement, and conduct of foster children *27th Yrbk. Nat Soc Stud Educ*, 1928, Pt 1, 103-217

10 GOODENOUGH, F L The Kuhlman-Binet tests for children of preschool age *Univ Minn. Instit. Child Welfare Monog Ser*, 1928, No. 2 Pp 146

11 GOODENOUGH, F L, & ANDERSON, J E Experimental child study New York Century, 1931. Pp xv+544

12 HAGGERTY, M E., & NASH, H. B Mental capacity of children and paternal occupation *J Educ. Psychol*, 1924, 15, 559-572

13. HALL, C F Compilation of the laws of Minnesota relating to children (General statutes 5356, 5365A, 5365C, 4554, 4550, 4553, 4573 4569, 4570, 4572, 4562, 4564, 4567, 8625, 4456, 4457, A B Opinion 3—Aug 19, 1918.) St. Paul, Minn. State Board of Control, 1918 Pp. 213

14 HART, H H Preventive treatment of neglected children New York. Charities Publ., 1910 Pp 419

15 JONES, D C, & CARR-SAUNDERS, A M. The relation between intelligence and social status among orphan children *Brit J Psychol*, 1927, 17, 343-364

16. KELLEY, T. L Statistical method New York Macmillan, 1923 Pp xi+390

17. PENDLETON, O New aims in adoption *Ann Amer Acad Pol. & Soc Sci*, 1930, 151, 154-161

18 SEARS, J B., PROCTOR, W., & WILLIAMS, J H Boise school survey. (*Educational Survey Series, No 9*) Yonkers, N Y. World Book Co Pp 290

19 SLINGERLAND, W. H Child placing in families New York Russell Sage Foundation, 1918 Pp. 260

20 SOROKIN, P, & ZIMMERMAN, C. C. Principles of rural-urban sociology New York Holt, 1929 Pp xv+662

21 TERMAN, L M, *et al* Genetic studies of genius Vol 1 Stanford University, Calif Stanford Univ Press, 1925 Pp. xv+648

22. UNITED STATES DEPARTMENT OF LABOR, Children's Bureau Bur Publ, No 75 Illegitimacy as a child welfare problem, Pt 2 (Emma O Lundberg & Katherine F Lenroot) 1921 Pp 408

23 YERKES, R M [Ed] Psychological examining in the United States Army *Mem Nat Acad. Sci*, 1921, 15 Pp vi+890

Institute of Child Welfare
University of Minnesota
Minneapolis, Minnesota

UNE ÉTUDE DE CERTAINS FACTEURS DE SÉLECTION LESQUELS INFLUENCENT LA PRÉDICTION DE L'ÉTAT MENTAL DES ENFANTS ADOPTÉS

(Résumé)

On a fait l'analyse de 6500 cas consécutifs de mères non mariées sur fiches au Bureau des Enfants de l'Administration de l'Etat de Minnesota, dans le but d'obtenir tous les renseignements possibles sur le choix des enfants illégitimes à placer dans des maisons d'adoption On a trouvé une vraie différence en faveur d'une meilleure éducation et d'une meilleure occupation pour les vraies mères des enfants adoptés que pour les vraies mères des enfants retenus On n'a pu dans aucun des deux cas expliquer la différence par l'âge ou demeure Une comparaison de l'occupation des grands-parents maternels et paternels montre aussi une différence qui favorise les grands-parents des enfants adoptés Ces données amènent à la conclusion que le milieu de famille des enfants illégitimes adoptés est supérieur à celui des enfants illégitimes en général, et la probabilité d'une hérédité productive d'une capacité moyenne ou meilleure est donc à espérer Quand on peut contrôler l'âge auquel les enfants sont placés dans les maisons d'adoption, on trouve que les mères des enfants placés à l'âge de 3 mois ou moins atteignent un niveau d'éducation constamment plus élevé que les mères dont les enfants sont placés à l'âge de 21 mois ou plus Quand on considère l'occupation des grands-parents, du père, et de la mère, on trouve que l'occupation des aieux est supérieure au cas des enfants placés l'âge de 3 mois ou moins Il résulte de cette évidence que, si l'état social et la capacité sont positivement associés les parents qui adoptent de très jeunes enfants obtiennent selon toute probabilité des enfants relativement supérieurs en capacité à d'autres enfants adoptés L'éducation et l'occupation des aieux ont donné une corrélation avec l'éducation et l'occupation des parents qui adoptent, d'environ 0,30 Donc dans toute étude de nature-nurture où il s'agit des enfants adoptés, il faut considerer la possibilité d'une ressemblance mentale entre le parent qui adopte et l'enfant tout-à-fait séparée de l'éducation et du milieu

LEAHY

EINE UNTERSUCHUNG GEWISSER SELEKTIVEN EINWIRKUNGEN IN DER VORAUSSAGUNG DES GEISTIGEN NIVEAUS BEI ADOPTIERTEN KINDERN

(Referat)

Man analisierte 6500 aufeinanderfolgende Protokolle unverheirateter Mutter die in dem Children's Bureau, State Board of Control, Minnesota, einregistriert waren, um moglichst viele Evidenz zu sammeln uber die Auslese illigitimer Kinder zur Unterbringung in Adoptivheimen Man fand einen zurverlassigen Unterschied zu Gunsten der eigentlichen Muttern adoptierter Kinder im Gegensatz zu den eigentlichen Muttern beibehaltener Kinder, indem die Mutter adoptierter Kinder statistisch eine hohere Leistung in Bezug auf Ausbildung und Beruf erwiesen Es zeigte sich eine gleich zuverlassiger Unterschied zu Gunsten der eigentlichen Vater adoptierten Kinder im Vergleich mit den eogentlichen Vatern beibehaltener (retained) Kinder In weder dem einen noch anderen Fall konnte

der Unterschied auf Alter oder Wohnung zuruckgefuhrt werden Eine Vergleichung der Berufstande der mutterlichen und vaterlichen Grosseltern offenbart ebenfalls einen Unterschied zu Gunsten der Grosseltern adoptierter Kinder Diese Befunde fuhren zu dem Schlusse, dass der Familienhintergrund adoptierter illigitimer Kinder dem von illigitimen Kindern im Allgemeinen uberlegen ist, und dass also die Wahrscheinlichkeit einer mittelmassige oder uberlegene Tuchtigkeit erzeigenden Hereditat besteht Kontrolliert man Alter zur Zeit der Unterbringung in Adoptivheimen, so zeigt es sich, dass die Mutter von Kindern, die mit 3 Monaten oder junger untergebracht worden sind, einen zuverlassig hoheren Erziehungsniveau erreichen, als Mutter, deren Kinder mit 21 Monaten oder spater untergebracht werden Es besteht Überlegenheit der Ahnen in Bezug auf Berufsleistung bei Kindern, die mit 3 Monaten oder fruher adoptiert worden sind, wenn man die Berufe der Grossvater, des Vaters und der Mutter berucksichtigt Aus diesen Befunden zieht man den Schluss, dass, wenn soziale Stellung und Tuchtigkeit positiv verbunden sind (are positively associated), so erhalten Adoptiveltern, die sehr junge Kinder nehmen, wahrscheinlich Kinder, die anderen adoptierten Kindern in Tuchtigkeit uberlegen sind. Ausbildung und Beruf der Ahnen lieferten einen Korrelationsziffer von ungefahr 30 mit Ausbildung und Beruf der Adoptiveltern Es muss also die Moglichkeit einer geistigen Ahnlichkeit zwischen Adoptiveltern und Kind, die von Erziehung und Umgebung vollstandig unabhangig ist, in jeder, adoptierte Kinder in Anspruch nehmenden, Untersuchung der Beziehungen zwischen Vererbund und Umgebung berucksichtigt werden.

LEAHY

A STUDY IN RATING TECHNIQUE WITH SPECIAL REFERENCE TO ACTIVITY IN PRESCHOOL CHILDREN*

From the University of Chicago

HELEN LOIS KOCH AND HELEN STREIT

INTRODUCTION

The usual method of rating a trait of personality involves the recording in some quantitative form of a general impression—an impression influenced probably to a greater or less degree by all of the rater's contacts with the subject being described. Ratings based upon these broad, more or less unanalyzed abstractions have, of course, gross limitations. Single events often weigh heavily in creating the rater's impression; abstraction frequently is not adequate, with the result perhaps that the halo from another trait colors the judgment on the quality under consideration; standards may shift with mood and immediate association, definition and analysis of trait indexes may be undiscriminating; limited contact of the rater with the subject may deprive the former of most pertinent information, etc.

The short-interval scheme of rating represents an effort to surmount some of the difficulties just enumerated. The essence of this method is the frequent observation of short samples of behavior and these widely scattered in time and circumstance. Various means of describing the behavior quality exhibited during the brief observation period have been employed. In some cases the mere presence or absence of an alleged objective index of the trait in question has been noted. If, for example, mouth manipulation is taken as a symbol of nervousness, the O may merely record whether or not there occurred any mouth play during an interval of ten minutes. Gradation or individual differences in the trait are supposedly to be revealed by the number of observation samples in which the trait index was observed. In still another scheme, the frequency of occurrence of the trait indicator in each interval is noted—the number of times

*Accepted by Carl Murchison of the Editorial Board and received in the Editorial Office, February 24, 1932

the child smiles in two minutes, for example These frequencies for the various intervals are then summed to obtain the final index.

Since the merit of the objective symptom is always a problem to be reckoned with, some workers have preferred to evaluate the behavior exhibited during each observation on the basis of all observable qualities—in other words, on the basis of the general impression or Gestalt While it is alleged that the latter method is lacking in the objectivity of the former, it has yet to be demonstrated that trained observers rating simultaneously, but independently, disagree radically in the impressions they record, if the interval observed is short enough—one-half minute, for instance

Another of the problems which confront investigators who use the short-time samples of behavior is that of determining the effect of the particular selection of observation intervals upon the returns Most of the studies heretofore reported have involved recordings made during a few hours of the day and have not systematically explored the influence of the situations which constitute the setting of the behavior This, of course, is to be regretted, for we may well question whether a child who is talkative when playing with his companions on the playground will be relatively talkative at a dinner table graced by the presence of a certain teacher Certainly it would be unwise to proceed far with the method without investigating the diagnostic significance of behavior samples obtained under each of a variety of circumstances. Perhaps even different standards for the trait, as it might be exhibited in different activities, need to be kept in mind

Objectives

The present study, which is one that is concerned largely with the technique of trait measurement, proposes to make inquiry into some of the problems just stated It purports (1) to adapt or devise a series of 11 rating scales for describing degrees of activeness in preschool children in different kinds of pursuits, (2) to test the objectivity of the ratings made in accordance with the scales by a comparative study of the simultaneous independent judgments of two observers, (3) to determine for a rather homogeneous group of preschool children the reliability of the returns from the use of each of 9 scales as well as for all combined, (4) to study the intercorrelations between the average ratings on activeness obtained as a result of the observation of performance in 9 different types of occupation,

(5) to assay the diagnostic import of performance in each department of activity; (6) to estimate the relative value of different scatterings and lengths of observation intervals, and (7) to discover whether level of activeness, or perhaps energeticness, is a function of mental age, chronological age, sex, time of day, temperature, and humidity

Subjects

The subjects of the study were 50 children in the University of Chicago Cooperative Nursery. They ranged in age from 30 to 62 months Eleven children, 30 to 35 months in age, were in the so-called two-year group; 11, 34 to 44 months, in the younger three-year group; 11, 47 to 51 months, in the older three-year group; and 17, 50 to 62 months, in the so-called four-year group

The Rating Scales

The rating scales used were, with one exception, adapted from those suggested by Sweeny, Hejinian, and Sholley (5). The adaptations consisted in combining scales, construing levels more specifically, and adding and dropping items A new scale—the one for dining—was devised. The gradations or steps in these scales were assigned arbitrarily the respective weights of 1 to 5

Each scale is independent of the others in its standards, just as would be an instrument draughted by a worker who planned to employ only one. The scaling, furthermore, while based on much careful observation has not been checked, as we had hoped it would be, by objective measures of energy output. Statistically, even, the grading is arbitrary In these respects, however, the scales are not at variance with prevailing ones

The classification of pursuits, performance in which our scales were devised to describe, while not entirely logical, is based largely upon the routines of the Nursery School Indoor and outdoor, free and directed, circumscribed and relatively unrestrained, preparatory and consummatory activities are classes represented in the scheme adopted The major consideration influencing the grouping, however, was the amount of body activity usually called out by the occupations We did not, though the question of the significance of different kinds of action is a pertinent one, feel justified in attempting to discriminate between voluntary and involuntary movement,

for example, or between what might be called directed and overflow activity

The scales are presented below

Indoor pedal activities

1 *a* Standing still or leaning against an object with little effort to maintain balance
   b Making slight movements

2 *a* Walking calmly or aimlessly for 10 seconds or less

3 *a* Walking with ordinary vigor from 10 to 20 seconds
   b Carrying a small object
   c Trying with moderate energy to open a door or drawer
   d Climbing a flight of stairs very slowly

4 *a* Running, hopping, or jumping for at least 10 consecutive steps or for 2 spurts of 5 steps each
   b. Carrying a chair or any other comparatively heavy object, such as an armful of blocks
   c Climbing a flight of steps at a moderate pace
   d Walking with ordinary vigor for over 20 seconds

5. *a* Running, hopping, jumping, climbing, or crawling in a strenuous manner
   b Carrying or dragging a heavy object, such as a table or a large basket of blocks

Floor play

1 *a*. Sitting quietly on the floor with little bodily movement

2. *a* Merely watching others while maintaining a squatting position
   b Sitting quietly, confining movements mostly to the arms
   c Shifting the trunk position twice or less

3 *a*. Moving about calmly (walking, creeping, or crawling) for practically the whole of the half-minute period
   b Remaining in a squatting or sitting position, but shifting position 3 times or more
   c Standing and stooping several times
   d Hammering or sawing for 5 strokes or less

4 *a* Moving about the room in lively pursuit of a game for 15 seconds or less
   b Hammering or sawing more than 5 strokes but not extremely forcefully
   c. Stooping frequently
   d. Picking up blocks energetically for 15 seconds or less

5 *a* Chasing, rolling, crawling, moving, etc, vigorously for 15 seconds or more
   b Hammering or sawing in a way that is not only vehement but also sustained for most of the period

Table and house play

1 *a* Sitting or standing quietly with few bodily movements but showing a passive interest in table or house activities

2 *a* Using calmly and in a mannerly fashion small equipment the manipulation of which calls for little effort
   b Sitting or standing at work

3 *a* Sitting quietly, in the main, but shifting the whole body or stooping once
  b Using the arms and hands rather actively
  c Moving the whole body once for a short distance in order perhaps to procure materials
4. *a* Moving from the table to the shelves or from the table to the floor twice
  b Shifting the gross body position or stooping twice
5 *a* Shifting or moving the whole body three times or more
  b Showing such energetic activity in table or in house play as pulling children about or pushing them away

Circle activities (the children are supposed to be sitting on chairs or on the floor)

1. *a*. Sitting quietly without interest
2 *a* Shifting position slightly two times or less
  b Sitting quietly but showing interest
  c. Maintaining an awkward position
3. *a*. Shifting position three times
  b. Participating in finger plays
  c. Wiggling considerably merely the hands
  d Singing alone before the group
  e. Keeping time gently to music with the hands
4 *a* Swinging the arms to music
  b Sprawling on the floor once
  c. Keeping time to music with the feet
  d Shifting position conspicuously four times
  e. Wiggling considerably with the whole body
5 *a* Lying on the back and keeping time with the legs
  b Tussling with the other children
  c. Using both hands and feet while keeping time to music
  d. Shifting position conspicuously more than four times

Circle activities (the children are engaged in rhythmic group exercises or games which involve moving about)

1 *a* Sitting alone or standing apart without participating in the exercises
2. *a* Walking about quietly
  b. Standing and using arm movements only
3. *a*. Walking actively
  b Wabbling along slowly and deliberately
  c Running slowly for 10 seconds or walking slowly for 10 seconds with movements inhibited voluntarily
4 *a*. Running for 10 seconds or more
  b. Participating in vigorous exercises from 10 to 20 seconds
5 *a*. Participating actively in vigorous exercises for over 20 seconds
  b Falling, rolling, rough-housing for most of the period

Dressing or undressing:

1 *a*. Sitting or standing with little movement
2 *a*. Dressing self, with or without help, in non-energetic manner
  b Working with small wraps (gloves, cap, etc) spasmodically
  c Walking about aimlessly
  d. Making unnecessary movements with the hands and feet

3 *a* Participating actively in work
b Working with small wraps continuously for the whole period
c Moving about rather actively but not attending to work
4 *a* Working 20 seconds with large wraps
b Showing unnecessarily vigorous activity for a short period—activity such as running, throwing, or dragging things about
5 *a* Working steadily on large wraps practically all of the 30-second period
b Jumping, falling off a bench, pushing, etc, for a period lasting 20 seconds or more

Resting·

1 *a* Resting quietly—no gross observable movements
b Making slight finger movements
2 *a*. Talking quietly to self
b Moving the arms or legs slightly three times or less
c Showing active interest in the other children by peering at them
3 *a* Moving the arms or legs more than three times
b Shifting the gross body position once
c Sitting up or resting on the arms quietly
4 *a* Moving about actively for 15 seconds or less
b Shifting the gross body position twice
c Kicking the feet
5 *a* Moving the body actively for 15 seconds or more
b Flouncing the blanket
c Getting up off the cot

Eating

1 *a* Sitting quietly, taking not more than 2 small bites of food
2 *a* Eating slowly, pouring milk, or chewing food
b Making slight movements with the hands or feet
3 *a* Eating slowly but continuously for 30 seconds
b Drinking steadily
c Taking several bites of food of fair size and chewing them
d Making many unnecessary but slight movements, or maintaining an unusual position
e Serving self or walking a few steps on legitimate business
4 *a*. Eating rapidly for 30 seconds
b Shifting the gross body position twice—getting off a chair
c Moving the trunk vigorously while eating
5 *a* Kicking, bending over, or showing unruly behavior in other ways in sustained fashion
b. Walking around the room vigorously

Outdoor pedal activities

1 *a* Standing quietly or leaning against an object other than a vehicle or a piece of apparatus
b Moving 2 or 3 steps
2 *a* Walking about calmly with slight accompanying activity
b Watching others while maintaining a squatting position
c Running 4 or 5 steps

3. *a* Walking energetically from 10 to 20 seconds
   b. Running calmly for 15 seconds or somewhat less
   c Walking slowly on skates
   d Walking slowly but with large steps
   e. Carrying a burden of moderate weight and size
4. *a*. Walking energetically from 20 to 30 seconds
   b Running rapidly for 15 seconds or less
   c Running calmly for 15 seconds or more
   d Skating
   e Jumping, scuffling, sliding on ice, etc, with moderate energy
5. *a*. Running rapidly for 15 seconds or more
   b Hopping, jumping, or tussling for 15 seconds or more

Apparatus and vehicle play

1. *a* Sitting with little bodily movement in or on a vehicle or a piece of apparatus
   b Being pushed gently on a swing or pulled in a wagon
2. *a* Riding a vehicle slowly
   b Riding energetically for spurts of 5 seconds or less with calm intervals between spurts
   c Using the apparatus spasmodically
   d Walking a low plank
3. *a* Manipulating vehicles actively
   b Pulling a moderately loaded wagon or pushing a wheelbarrow
   c. Pushing someone in a swing gently
   d Climbing, swinging, teetering, etc, moderately actively
4. *a* Riding or using equipment energetically for a considerable period
   b Pulling a heavy load or pushing a loaded wheelbarrow up an incline
   c Teetering, pumping on a swing, climbing, or jumping vigorously
5. *a*. Riding rapidly or using equipment very energetically
   b Pulling a heavy load for 15 seconds or more
   c. Riding over difficult surfaces for 15 seconds or more
   d Using the apparatus in an unusually strenuous manner

Construction activities on the playground.

1. *a*. Sitting on the ground, sand, or something other than a vehicle or a piece of apparatus
2. *a* Using light materials (spoons, small buckets, etc) spasmodically
   b Moving the body very little
   c. Playing house quietly
3. *a*. Using light materials actively and continuously
   b Using the heavier materials (large shovels, large blocks, etc) spasmodically
   c Walking rapidly for 10 seconds, or slowly for 20 seconds, as a part of the play construction activities
4. *a* Using the heavier materials continuously
   b. Stooping four or more times
   c Walking rapidly for more than 10 seconds when it is part of house play
   d. Lifting or dragging heavy objects for 15 seconds or less
5. *a*. Lifting or dragging heavy objects for more than 15 seconds

Observation Procedure

The program of long-sustained observation employed by Sweeny, Hejinian, and Sholley (5) was abandoned entirely and a short-interval observation scheme substituted. Each child was observed only a minute at a time in each occupation, two ratings being given during this minute, one on the activity level of the first half-minute, one, on the second. It was our intent to obtain ratings on 30 one-minute samples of each of 50 children's behavior in each of at least nine activities. For the 20 children who remained at the nursery during the afternoon session, samples were to be taken during 14 different activities.

In our effort to obtain 60 half-minute samples we were only seldom successful. More time was consumed with the procedures than was anticipated, the session terminated, and the children left for vacations before the observations could be completed. Absences and illnesses thwarted us some, but a greater difficulty was met in the fact that some children practically never engaged in a certain type of pursuit. Some, for example, never played on the floor, some, never with materials at a table.

For the nine types of employment occupying the morning hours, a range of 7 to 49 half-minute observations for a child in an activity was our accomplishment, whereas for the five occupations occurring in the afternoon, the range was 20 to 82. Ninety-three per cent of the 500 averages for the morning activities are based upon 25 or more recordings, 98 per cent, on 20 or more. For the afternoon series all but 2 of the 140 averages are computed on the basis of 30 or more recordings.

Observations of the various nursery groups we rotated, a given group being under scrutiny in all of its activities on a given day. Within a group, the observations of the children were also rotated. Deviations from a straight-forward sequential scheme were common, however. Some children were absent much. In order to accumulate ratings on these S's, observation was continued later into the session by several weeks than for the rest of the group. Occasionally, even, judgments were made during two-minute observation periods on the same day for these children. When the latter was done, the ratings were separated by as large a time interval as arrangements permitted. We were also inclined, in the case of certain subjects who rarely engaged in a particular kind of activity,

to rate them as soon as we saw them involved in the occupation. Even with this compromise we did not in one case succeed in obtaining more than 10 ratings Occasionally, too, the rotation order was sacrificed in order to rate all of the children during a period which was being curtailed for some reason by the nursery teacher.

While, then, our success in adhering to a certain fixed rotation scheme was far from perfect, the basis for selecting a particular child for observation at a particular time, was, with one exception—the case of the individuals who practically never interested themselves in a pursuit—not what the child was doing but something wholly unrelated to his performance Hence, the biasing of our results through the observation of conspicuous or unusual behavior is not likely.

In general, we essayed, as has been stated, to appraise on the same day the performance of all of the children in a given play group in all of the activities. Since there were four such groups, we succeeded in observing each child about once a week for one minute in each activity. Such factors as temperature, humidity, and the peculiar features of a particular day were, therefore, not constant from group to group These factors were similar only to the degree that a week's range mediates uniformity

All of the foregoing compromises with unavoidable irregularities should be kept in mind in interpreting our reliability coefficients, for the former would undoubtedly operate to lower the latter and, hence, mask the possibilities of the general method. Average ratings on only one activity, appraised with such factors controlled as we found it necessary to vary in order to make observations on the same day on a considerable number of occupations, would yield reliabilities much higher than ours, which might be viewed almost as minima.

Objectivity of the Ratings

The possibility of developing a rating method which would be reasonably objective we attempted to determine by having two O's for 30 days simultaneously, though independently, rate the children in the nine types of occupation that could be observed at the Nursery School in the mornings For the five initial days no records were kept. In Table 1 is given for the next five weeks for which we have records the percentage of judgments which were in agreement By the fifth week of the training, exact correspondence in 90 per

TABLE 1

PERCENTAGE OF AGREEMENT BY WEEKS BETWEEN THE SIMULTANEOUS RATINGS MADE BY TWO OBSERVERS ON THE VARIOUS ACTIVITIES

Activity	Week					
	First	Second	Third	Fourth	Fifth	Average
Pedal (indoors)	68.0	75.0	77.4	70.4	81.7	74.5
Floor	67.7	81.4	78.6	80.3	90.5	79.7
Table	73.3	71.9	73.2	74.4	73.1	73.2
Circle (sitting)	50.6	75.0	78.0	71.6	78.2	70.7
Circle (standing)	26.9	69.0	50.0	79.8	72.0	59.5
Dressing-undressing	*60.0*	*56.6*	*69.8*	*66.1*	85.0	*67.5*
Pedal (outdoors)	68.9	78.0	74.3	87.0	81.1	77.9
Apparatus	55.3	70.8	84.9	75.5	92.2	75.7
Construction	66.7	65.8	73.8	67.7	76.7	70.1
Total	59.0	71.3	74.6	73.8	81.1	72.0

TABLE 2

INTER-OBSERVER CORRELATIONS BASED ON SIMULTANEOUS RATINGS MADE DURING A PRACTICE PERIOD FOR FIVE WEEKS

Activity	r	PE_r
Pedal (indoors)	.97	.01
Floor	.90	.02
Table	.85	.02
Circle (sitting)	.85	.02
Circle (standing)	.84	.03
Dressing-undressing	.92	.01
Pedal (outdoors)	.93	.01
Apparatus	*.89*	*.02*
Construction	.90	.02

cent or more of the valuations was attained for two of the occupations; and in 75 per cent or more, for the other seven. The average percentage of judgments in agreement for all activities in this fifth week was 81. It is worthy of comment that seldom were the ratings at variance by more than one point in the scale.

The correlations between the averages of the scores given by the two O's during these early weeks are presented in Table 2. Since the number of observations for a child in each activity was usually below 10 and the ratings were products of all stages of the training period, these correlations are encouragingly high.

Discrimination of gradations in the quieter activities—i.e., activities having a more restricted range than did the floor and apparatus

play—were accomplished with greatest difficulty The major embarrassment encountered, however, was that of keeping in mind at the same time the standards for the different scales which were being used parallel Since, unfortunately, we desired to rate a child on all activities on a given day, the difficulty seemed an unavoidable one and one to be surmounted only by practice. In the case of any one scale employed singly, we should expect a high degree of agreement between the observers to be reached within a period of about two weeks.

It seems, then, that personal bias need not be a very important factor in the coloring of ratings made on short samples of certain forms of behavior, especially if rather arbitrary definition or legislation is employed with respect to significant controversial details. We would insist, however, that while it is apparent that the scales used do show some arbitrary specification of detail, the ratings, after all, were based largely upon the total Gestalt, the intensity, extensity, and duration of the activity always being thrown into the balance

Reliability of the Ratings

The reliability of our measures we attempted to evaluate in two ways—one through correlating the two sets of mean valuations resulting from averaging the judgments made in alternate minutes, the other, through correlating the rating averages based upon the estimates made during the first and second halves, respectively, of the minute observation period. We have had in mind, for instance, such questions as the number of minute samples of behavior with the shifts in mood, health, etc., that a week effects (the children were observed in each activity only about once a week), that will be necessary for highly reliable measures. We were also curious to know whether the activity level was so well sustained over a minute as to make 30 seconds of observation about as illuminating as 60 seconds

Table 3 gives the reliability coefficients. For the alternate-minute scheme, the corrected coefficients range from +.44 to +.75, with 6 out of the 9 being +.70 or above If one considers the number of factors we have already listed which would tend to reduce reliability, considerable faith in the general method seems justified If correlations as high as those obtained are the result of approximately 15 minutes' observation, for most of the activities one hour to an hour and a half of attention scattered over 120 to 180 successive days, for

TABLE 3

RELIABILITY COEFFICIENTS FOR THE VARIOUS ACTIVITY RATINGS

Activity	Alternate-minute scheme (unweighted scores)	First- vs second-half-minute scheme (unweighted scores)	Alternate-minute scheme (weighted scores)	Spearman-Brown correction for length—alternate-minute scheme (unweighted scores)	Spearman-Brown correction for length—first- vs. second-half minute scheme (unweighted scores)	Spearman-Brown correction for length—alternate-minute scheme (weighted scores)
Pedal (indoors)	+ 59± 06	+ 58±.06	+ 57± 06	+.74± 05	+ 73±.05	+ 72± 05
Floor	48± 07	67± 05	53± 07	65± 07	.80± 04	69± 06
Table	54± 07	81± 03	57± 06	70± 06	89± 02	72± 05
Circle (sitting)	59± 06	77± 04	60± 06	74± 05	87± 02	75± 05
Circle (standing)	57± 06	.73± 04	60± 06	72± 05	84± 03	75± 05
Dressing-undressing	60± 06	81± 03	61± 06	75±.05	.89± 02	.76± 05
Pedal (outdoors)	57± 06	76± 04	60± 06	.74± 05	86± 03	75± 05
Apparatus	44± 08	72± 05	43± 08	61± 08	84± 03	60± 08
Construction	28± 09	92± 01	32± 09	44± 11	96± 01	48± 10
Total	78± 04	91± 02	83± 03	88± 02	95± 01	91± 02

example, should furnish a measure adequate for individual diagnosis in a group showing the degree of homogeneity that ours did

The correlations between scores based upon valuations made in successive half-minutes ranged from +.73 to +.96, 7 of the 9 coefficients falling above +.80. This seems to indicate that an activity level tends to be fairly well sustained over a minute and chance errors correlated—facts which, in turn, suggest that the same amount of time divided into 60-second observation periods would yield a less satisfactory index than if divided into 30-second periods and scattered accordingly over double the span of time.

A conspicuous exception to the foregoing observation is the indoor pedal activity Here, samples picked within the same minute seemed to yield as much variety as samples taken from alternate weeks

The averages of the ratings on outdoor construction activities manifest the least reliability; while those on dressing and undressing exhibit the highest The small coefficient in the former case is probably due to the fact that several children practically never constructed anything while on the playground. We have, consequently, few samples of their behavior and correspondingly great irregularity in result. In support of this hypothesis, we would call attention to the high coefficients obtained from correlating scores based on successive half-minutes

The fact that the reliability coefficients for dressing and pedal activities are as high as they are may be attributable, in part at least, to the fact that the scope of the activity patterns which these occupations permit or suggest is somewhat more circumscribed than is true for the other pursuits. In accordance with this hypothesis we note that apparatus and floor play, where we should expect the most variation in pattern, do show the lowest reliabilities

Mistrusting the weightings represented by our simple arbitrary quantitative judgments of 1 to 5, we, assuming that activity distributes itself according to the normal curve, weighted our scale values in conformity with their frequency among our ratings. The formula used was $\frac{z_1 - z_2}{\frac{f_2}{n}}$ (4, p 221) Computing the reliabilities on the basis of these weighted scores resulted in coefficients so little different in most cases from those obtained with unweighted scores, it has seemed justifiable for the sake of simplicity in computation to pro-

ceed with the latter (See Table 3.) We might add that the weighted and unweighted scores correlate to the extent of +.99 in 6 cases and above +.96 in all cases

INTERACTIVITY CORRELATIONS

Of fundamental concern to us is the question of the variation in the relative activity of an individual from occupation to occupation Can, for example, estimates of activeness descriptive of performance in one type of activity be expected to furnish a satisfactory delineation of the behavior quality to be observed in other pursuits. Table 4 presents our answer to this question. The interactivity correlations are low and in a few cases slightly, but probably not significantly, negative. Even when the raw coefficients are corrected for attenuation, about two-thirds of the correlations are below + 50. Pedal activities, indoor and outdoor, correlate highest with each other—+ 52 when uncorrected and + 71 when corrected for attenuation. The ratings on play involving building with out-of-door materials correlate significantly with the ratings on energeticness of locomotion in the out-of-doors The activity levels displayed in indoor pedal and dressing undertakings seem also to be related to a small degree. Among the unusually low *r*'s—*r*'s under + 10—are those representing the relation between the levels shown in apparatus and table work, pedal (outdoors) and circle (standing), circle (sitting) and dressing-undressing, circle (standing) and table, construction and circle (standing), and circle (standing) and circle (sitting). The small coefficients, it seems unlikely, are entirely the result of arbitrary scaling or inaccurate estimates

It is interesting that valuations made during the circle activities, which are characterized by much spirit and snap, correlate so little with assessments made during most of the other occupations Since in the former the child is under direction and is being stimulated to marked activity by the teacher's commands as well as by the example of the other children, it is probable his more usual tendencies have less than the normal opportunity to manifest themselves

Although the correlations for the afternoon activities are based on so few cases, they are worthy of some comment, the number of observations for a subject being actually considerably greater than for the morning activities. The latter factor is probably responsible for the higher interactivity *r*'s obtaining between situations corre-

TABLE 4

INTERACTIVITY CORRELATIONS—UNCORRECTED AND CORRECTED FOR ATTENUATION*

Activity	Pedal (indoors)	Floor	Table	Circle (sitting)	Circle (standing)	Dressing and undressing	Pedal (out-doors)	Apparatus	Construction	Total
Pedal (indoors)		$+36\pm08$	$+20\pm09$	$+36\pm08$	$+31\pm09$	$+43\pm.08$	$+52\pm07$	$+37\pm08$	$+22\pm09$	$+74\pm04$
Floor	$+52\pm12$		$+31\pm.09$	$+21\pm09$	$+21\pm09$	$+41\pm08$	$+22\pm09$	$+35\pm08$	$+13\pm09$	$+.63\pm06$
Table	$+28\pm13$	$+46\pm12$		$+35\pm08$	-06 ± 10	$+12\pm09$	$+20\pm09$	$+07\pm10$	$+17\pm09$	$+43\pm08$
Circle (sitting)	$+49\pm11$	$+30\pm13$	$+49\pm11$		-14 ± 09	$+06\pm09$	$+22\pm09$	$+.18\pm09$	$+21\pm09$	$+47\pm07$
Circle (standing)	$+42\pm11$	$+31\pm13$	-08 ± 01	-19 ± 13		$+34\pm08$	$+.03\pm10$	$+28\pm09$	-10 ± 09	$+42\pm08$
Dressing-undressing	$+58\pm10$	$+59\pm11$	$+17\pm01$	$+08\pm01$	$+47\pm11$		$+.35\pm08$	$+32\pm.09$	$+14\pm09$	$+65\pm06$
Pedal (out-doors)	$+71\pm09$	$+32\pm.13$	$+28\pm13$	$+30\pm12$	$+04\pm14$	$+48\pm11$		$+35\pm08$	$+49\pm.07$	$+64\pm06$
Apparatus	$+55\pm12$	$+.56\pm13$	$+11\pm02$	$+.27\pm.14$	$+42\pm13$	$+47\pm12$	$+53\pm12$		$+20\pm09$	$+.60\pm06$
Construction	$+39\pm16$	$+.25\pm18$	$+30\pm16$	$+37\pm16$	-18 ± 02	$+25\pm16$	$+88\pm12$	$+38\pm17$		$+.50\pm.07$
Total	$+91\pm04$	$+83\pm07$	$+55\pm.09$	$+58\pm09$	$+53\pm09$	$+80\pm06$	$+80\pm06$	$+82\pm07$	$+81\pm11$	

*The correlation coefficients above the diagonal are uncorrected, the coefficients below the diagonal are corrected for attenuation.

TABLE 5
INTERACTIVITY CORRELATIONS FOR AFTERNOON ACTIVITIES*
(20 cases)

Activity	Pedal (out-doors)	Apparatus	Construction	Dressing and undressing	Meal
Pedal (outdoors)					
Apparatus	+75± 07				
Construction	+85±.04	+63± 10			
Dressing-undressing	+53± 11	+63± 10	+44± 13		
Meal	+09± 16	+06± 16	+08± 16	−12± 16	
Rest	−18± 15	−28± 15	−22± 15	−14± 15	+21± 15

*The Spearman rank-difference method was used to compute these correlations and those in the following table because of the paucity of cases.

sponding presumably to some observed in the morning (See Table 5) It is necessary to remember, however, that the group involved in the two measures is not the same, the afternoon group being relatively more heterogeneous in age and interests than the total morning group.

In the face of the fair size of the r's between pedal, apparatus, construction, and dressing activities, the low and even negative correlations obtaining between the indexes of performance in the aforementioned pursuits and those for performance at meals or at the nap period stand out in strong relief. Though the magnitude of the probable errors of these latter coefficients make the minus character they occasionally show questionable, the r's do lead one to suspect that the child who moves much during the rest period may not be the one who plays and flits about most vigorously at other times. Perhaps the motility at nap time is colored strongly by action of a more or less involuntary or overflow nature It seems reasonable, too, that the frail child without an appetite will be fidgety at meals.

Still another of the interesting relations that Tables 5 and 6 reveal is that some of the interactivity correlations among dissimilar pursuits occurring in the afternoon are higher than those obtaining for similar occupations observed both in the morning and afternoon One of the latter is even negative

These generally low interactivity correlations, then, indicate, among other things, that the halo effect, which so mars gross general ratings, does not have a very perceptible influence on the returns

TABLE 6

CORRELATIONS BETWEEN RATING AVERAGES BASED ON SIMILAR MORNING AND AFTERNOON ACTIVITIES

Activity	ρ	PE_ρ
Dressing-undressing	+.72	05
Pedal (outdoors)	+.52	07
Apparatus	—21	10
Construction	+ 16	10

from the short-sample method. The correlations seem to suggest also that activeness is a function, in part at least, of the particular situation in which the child is placed. And index of activeness which would be accurately descriptive, should, hence, take into account the full gamut of the child's program during the day. Even our measures based upon morning activities at the Nursery School are probably rather inadequate. The somewhat low correlations of similar morning with afternoon activities furnish a line of evidence favoring this thesis.

DIAGNOSTIC OR PROGNOSTIC IMPORT OF SPECIFIC SERIES OF RATINGS

If we checked each index against the best total we have—i.e., that based on the means of all the morning activities—the order of predictive effectiveness would be as follows: pedal (indoors), dressing-undressing, pedal (outdoors), floor play, apparatus, construction, circle (sitting), table and circle (standing).

The correlations of each of the means for the various types of occupation with the average of all of these means range from +.42 to +.74. When, however, the coefficient of correlation between intra-mural pedal activities and the total is corrected for attenuation, it is as high as +.91. Floor, dressing, pedal (outdoors), apparatus, and constructive activity scores give corrected coefficients of correlation with the total of +.80 or better. These r's, of course, are spuriously high because the measure correlated with the average is included in the average.

It seems apparent from the foregoing findings that activeness as we judged it in the more circumscribed and directed activities is not so good an indicator of general activity level as our estimates made during those occupations permitting much freedom. We are in-

clined to admit, moreover, that our whole method is such as to limit seriously in the final result the effect of the individual's choice of occupation A sampling scheme such as Goodenough (3) employed, which would take cognizance of choice of task or amusement, has much to commend it It is possible, for example, that a child who plays strenuously on the apparatus almost all day long will, when he settles down to table work, be no more energetic than the child who spends most of his day doing handwork

Partial Correlations

The nature of the factors mediating such degrees of interactivity correlation as we observed is a question which might with profit be raised. Are these due to relatively general factors, such as we might suspect to be present in maturity, brightness, energeticness, or even nervousness or to many small, common, overlapping, specific elements in the adjustments demanded

That the latter is a partial answer we have already suggested In the constructive activities on the playground, for example, the children walk about much, seeking for materials wherewith to construct, dragging supplies to the proper corner, transporting tools, etc In the case of the relation between dressing and pedal activities, however, the hypothesis of many highly specific common elements does not seem quite so cogent, though it is still very plausible

The relationship to chronological age of activeness in each of

TABLE 7

Zero-Order and Partial Correlations between Activity Ratings and Chronological and Mental Age

1 = Activity, 3 = CA, 4 = MA

Activity	r_{13}	r_{14}	r_{34}	$r_{13.4}$	$r_{14.3}$
Pedal (indoors)	+ 36± 08	+ 34± 08		+ 15± 09	+ 07± 09
Floor	+ 28± 09	+ 28± 09		+ 09± 09	+ 09± 09
Table	— 10± 09	— 19±.09		+ 11± 09	— 20± 09
Circle (sitting)	— 07± 10	— 02± 09		— 10± 09	+ 07± 09
Circle (standing)	+ 29±.09	+.36± 09		— 02± 10	+ 22± 09
Dressing-undressing	*+ 63± 06*	*+ 62±.06*		+ 26± 09	+ 22± 09
Pedal (outdoors)	+ 07± 10	+ 13± 09		— 07± 09	+ 13± 09
Apparatus	+ 20±.09	+ 32± 09		— 13± 09	+ 29± 09
Construction	+ 05± 10	— 02± 09		+ 12± 09	— 11± 09
Total	+ 32± 09	+ 35±.08	+ 84±.03	+ 05± 10	+ 16±.09

the occupations is suggested by the data included in Table 7. The range of r's is from —07 to +.63. It appears that the older children dress and undress significantly more energetically than do the younger who, perhaps discouraged by the size of the task, are passive before it. The only occupations, activeness in which has the appearance of being unquestionably though slightly related to age, are floor play, indoor pedal, circle (standing), and dressing activities. Hence, we have probably the reason for the fact that the total activity score correlates only to the extent of + 32 with age.

The relation between mental age, as measured by the Kuhlmann-Binet test, and the activity ratings shows to be about the same in magnitude and direction as that noted for chronological age. Apparatus play manifests a slightly, but probably not significantly, greater relation to mental than to chronological age, whereas constructive plays shows less. On the whole, it seems safe to say that with the exception of dressing, the relationship of the activity scores to mental age is either small or negligible The r between the mean of all of the occupation means and mental age is no more than +.35.

Holding mental age constant does reduce somewhat the correlation between all but two of the activity ratings and chronological age. The r's between table and construction activities and chronological age are raised, the former being altered in sign as well In other words, if all the children were the same mental age, the oldest or more mature would tend to be to a small extent, if at all, more energetic only in the dressing-undressing activities. A slight negative, but again probably not significant, relation with maturity is observed for both forms of circle, outdoor pedal, and apparatus activities.

Partialling out chronological age from the relation between mental age and our measures alters most noticeably the r's between the former and pedal (indoors), floor, circle (sitting), dressing-undressing, and construction activities. The data of Table 7 suggest that, if all S's were the same chronological age, the mentally oldest, and hence the brightest, would tend slightly to be the more energetic in the circle (standing), dressing, and apparatus activities; whereas the duller would tend slightly to be the more active in table play

The interactivity correlations in the production of which chronological age is a factor, to the extent that partialling out CA reduces by as much as 9 points or more (about 1 PE) the r's between

TABLE 8
INTERACTIVITY CORRELATIONS—CHRONOLOGICAL AGE AND MENTAL AGE PARTIALLED OUT

Activities correlated	Zero-order correlations	Chronological age partialled out	Mental age partialled out	Chronological and mental age partialled out
Pedal (indoors) and floor	+ 36± 08	+ 29± 09	+ 29± 09	+ 29± 09
Pedal (indoors) and table	20± 09	.25± 09	29± 09	27± 09
Pedal (indoors) and circle (sitting)	36± 08	.41± 08	39± 08	41± 08
Pedal (indoors) and circle (standing)	31± 08	23± 09	21± 09	22± 09
Pedal (indoors) and dressing-undressing	43± 08	28± 09	30± 09	27± 09
Pedal (indoors) and pedal (outdoors)	52± 07	53± 07	51± 07	53± 07
Pedal (indoors) and apparatus	37± 08	33±.08	29± 09	32±.09
Pedal (indoors) and construction	22± 09	22± 09	24± 09	23± 09
Pedal (indoors) and total	74± 04	71± 05	70±.05	71±.05
Floor and table	31± 09	30±.09	.27± 09	33± 08
Floor and circle (sitting)	21± 09	20± 09	22±.09	19± 09
Floor and circle (standing)	21±.09	14± 09	12± 09	12±.09
Floor and dressing-undressing	41± 08	31± 09	.31± 09	.30± 09
Floor and pedal (outdoors)	22± 09	21± 09	.19± 09	.20± 09
Floor and apparatus	35± 08	31± 09	29± 09	30± 09
Floor and construction	13± 09	12± 09	.14± 09	13± 09
Floor and total	.63± 06	59±.06	59± 06	.59± 06
Table and circle (sitting)	35± 08	35± 08	35± 08	37± 08
Table and circle (standing)	— 06±.10	— 03± 10	09± 09	01± 10
Table and dressing-undressing	+ 12± 09	+ 24± 09	31± 09	30± 09
Table and pedal (outdoors)	.20± 09	.27± 09	23± 09	.30± 09
Table and apparatus	.07± 09	12± 09	14+ 09	19+ 09
Table and construction	17± 09	23± 09	17±.09	21±.09
Table and total	43± 08	63± 06	54±.07	68±.05
Circle (sitting) and circle (standing)	— 14± 09	— 13± 09	—.14± 09	— 15± 09
Circle (sitting) and dressing-undressing	+ 06± 10	+ 13± 09	+ 09± 09	+ 12± 09
Circle (sitting) and pedal (outdoors)	22±.09	23± 09	22± 09	22± 09
Circle (sitting) and apparatus	18± 09	20± 09	20± 09	19± 09
Circle (sitting) and construction	21± 09	21± 09	21± 09	22± 09
Circle (sitting) and total	47± 07	.52± 07	51± 07	52± 07

TABLE 8 (*Continued*)

INTERACTIVITY CORRELATIONS—CHRONOLOGICAL AGE AND MENTAL AGE PARTIALLED OUT

Activities correlated	Zero-order correlations	Chronological age partialled out	Mental age partialled out	Chronological and mental age partialled out
Circle (standing) and dressing-undressing	34± 08	21± 09	.16± 09	17± 09
Circle (standing) and pedal (outdoors)	03± 10	01± 10	— 02± 10	— 02± 10
Circle (standing) and apparatus	.28± 09	.24± 09	+ 18± 09	+ 19± 09
Circle (standing) and construction	— 10± 09	— 12± 09	— 10± 09	— 10± 09
Circle (standing) and total	+ 42±.08	+.36± 08	+ 34± 08	+ 34± 08
Dressing-undressing and pedal (outdoors)	35±.08	39± 08	35± 08	37±.08
Dressing-undressing and apparatus	.32± 09	25± 09	16± 09	20± 09
Dressing-undressing and construction	14±.09	14± 09	16± 09	17± 09
Dressing-undressing and total	.65± 05	61± 06	59± 07	60±.06
Pedal (outdoors) and apparatus	35± 08	.34± 08	33± 08	32± 09
Pedal (outdoors) and construction	49± 09	49± 07	.50± 07	51± 07
Pedal (outdoors) and total	64± 06	65± 05	64± 06	.64± 06
Apparatus and construction	.20±.09	19± 09	22± 09	23± 09
Apparatus and total	.60± 06	.58±.06	55± 07	56± 07
Construction and total	.50± 07	.51± 07	54±.07	54± 07

them, are: pedal (indoors) and dressing, dressing and floor, and circle (standing) and dressing Holding chronological age constant increases the correlation between table and dressing, table and pedal (indoors), and table and the mean of all activities.

Whatever is represented by the Kuhlmann-Binet mental age is a substantial factor in causing the correlation between pedal (indoors) and circle (standing), pedal (indoors) and dressing, circle (standing) and dressing, and circle (standing) and apparatus In four cases—pedal (indoors) and table, table and dressing, table and

apparatus, and table and total—the interactivity correlation was increased somewhat by partialling out mental age.

It is noteworthy, then, that neither chronological nor mental age is strongly influential in fully half of the interactivity correlations; and when the former are held constant, they still leave correlations of significant size in many cases.

Since chronological and mental age contribute not all that is responsible for most of the interactivity correlations, we may well look to other rather general or overlapping specific factors. The method of tetrad differences not being feasible with our data, as far as locating a general factor is concerned, it would be difficult to tell whether relatively general or more specific influences determine the pattern of our findings, and still more difficult to label these factors.

Sex Differences

A few variables not intimately related to questions of technique but which still may be factors influencing the functions we were observing, it has seemed well to investigate. Table 9 suggests that the sample of boys we dealt with is more active than the sample of girls. Since our age sex groups have so limited a membership, no detailed statistical treatment of the data seemed warranted. It is only the appearance in all four age groups of the finding favorable to the boys—if a high activity score can be viewed as favorable—that makes us respect it. The trend, too, is consonant with results reported by authors, among which are Bridges (1), Goodenough (3), and Farwell (2).

Whether the fact that the sex difference which our findings reveal decreases with age is to be taken seriously is a question. The

TABLE 9

Total Activity Rating of Age and Sex Groups

	Boys		Girls		Total	
Group	Number of *cases*	Activity *rating*	Number of *cases*	Activity *rating*	Number of *cases*	Activity *rating*
Two-year group	4	2.91	7	2.69	11	2.77
Young three-year group	7	2.93	4	2.76	11	2.87
Older three-year group	6	2.96	5	2.90	11	2.93
Four-year group	7	2.95	10	2.93	17	2.94

boys in the different age groups score about the same, but the girls present higher activity scores with each succeeding age. This is contrary to what Goodenough (3) reports Perhaps the Nursery School, with its regime of freedom, is counteracting in a measure the taboo or tradition favoring feminine inactivity.

Time of Day

Having the opportunity to observe 20 of the children in similar pursuits in the morning and afternoon, we have asked ourselves whether, even though the children had had a nap, they were any less vigorous in the post-meridian than in the *ante-meridian period.* Of course, we must remember that the social relationship in this reduced group of 20 which remained at the nursery after lunch may be very different from those prevailing when the 50 children are around as they are in the morning. Whatever the cause, then, three of the comparisons permitted by Table 10 indicate that activity was greater in the forenoon, construction being the only activity presenting a trend not in entire conformity with the others

Temperature

Since the temperature that was maintained in the different rooms of the nursery housing the various age groups presented not a great amount of overlapping, and since age is apparently a factor affecting activity slightly, we averaged for each age group separately the ratings given when the temperature was above and when below the median for the particular group. Nineteen out of the 25 comparisons made possible by Table 11 show the lower activity scores to be associated with the lower temperature. It is, of course, to be remembered that the temperature range considered was not impressive and few excursions occurred beyond the sixties and seventies.

TABLE 10
Activity Ratings for Morning and Afternoon
(20 cases)

Activity	Morning	Afternoon
Pedal (*outdoors*)	2 84	2 67
Apparatus	2 79	2 69
Construction	3 05	3.13
Dressing-undressing	2.88	2 66
Total	2 90	2 79

TABLE 11

ACTIVITY RATINGS FOR VARIOUS TEMPERATURE CONDITIONS INDOORS

Group	Two-year group		Young three-year group		Older three-year group		Four-year group		Total	
Average of *temperatures* above and below median	69°	73°	72°	78°	69°	74°	68°	74°	69°	74°
Pedal indoors)	2.81	2.86	2.99	3.05	3.05	3.12	3.19	3.21	3.04	3.11
Floor	2.91	2.94	3.04	2.94	3.11	3.15	3.01	2.97	3.03	3.01
Table	2.84	3.13	3.05	2.97	3.28	3.20	2.80	2.91	2.97	3.03
Circle (sitting)	2.51	2.58	2.87	2.94	2.73	2.88	2.58	2.68	2.61	2.72
Circle (standing)	3.01	2.99	2.51	2.60	2.83	2.94	3.05	3.19	2.99	3.06
Total	2.80	2.90	2.88	2.89	3.02	3.04	2.94	2.95	2.93	2.97
Average	2.82	2.90	2.89	2.90	3.00	3.06	2.93	2.99	2.93	2.99

The data for the pursuits engaged in out-of-doors, where the temperature contrasts were much greater and temperature average generally much lower than that prevailing indoors, show no very consistent trends If anything is true, it is that the lower temperature is associated with the greater activity (See Table 12.) Since we recall that a very heavy and lasting snow, which was, in turn, accompanied by cold weather, was the occasion of much vigorous play, we are inclined to place little faith in the hypothesis that temperature as such is responsible for the relationships suggested by our data.

TABLE 12

ACTIVITY RATINGS FOR VARIOUS TEMPERATURE CONDITIONS OUT-OF-DOORS

Time of day	A.M.		P.M.	
Average of temperatures above and below median	35°	55°	37°	58°
Pedal	2.73	2.66	2.61	2.74
Apparatus	2.75	2.79	2.72	2.66
Construction	3.12	2.96	3.30	3.08
Total	2.84	2.81	2.82	2.85
Average	2.87	2.80	2.86	2.83

TABLE 13

ACTIVITY RATINGS FOR VARIOUS CONDITIONS OF HUMIDITY OUT-OF-DOORS

Time of day	A.M.		P.M.	
Average of humidity recordings above and below median—%	49	79	44	73
Pedal	2.71	2.67	2.68	2.66
Apparatus	2.77	2.77	2.68	2.70
Construction	3.01	3.02	3.15	3.19
Total	2.83	2.81	2.83	2.84
Average	2.83	2.82	2.84	2.85

HUMIDITY

The indoor humidity varying very little, it did not seem profitable to comb our data for evidence of any relationship between hygrometer readings and activity. Our findings furnish, moreover, no convincing evidence that activeness is a function of humidity in the rather considerable extra-mural range observed.

SUMMARY

1. The activeness of each of 50 preschool children was observed for about 15 minutes in each of nine occupations occurring during the morning session at a nursery. The observation time was divided into one-minute periods which were scattered at intervals of about one week. Two ratings were made in each minute in accordance with the appropriate one of the nine five-point ratings schemes devised. For the 20 of the 50 children who were available for study in the after-, as well as in the forenoons, similar attention was given their behavior in five additional types of pursuits.

2. Two experimenters making judgments upon the same behavior independently reached in six weeks of training exact agreement in 90 per cent or more of their valuations for two of the occupations and in 75 per cent or more for all of the other activities. A number of the correlations between the means of the ratings in each pursuit given by the two observers during the whole of the training period are +.90 or above, one being as high as +.97. This fact indicates that the method can be made reasonably "objective."

3. The corrected reliability coefficients for the series of assessments made on each of the various occupations range from +.44 to +.74, when the items correlated were the means of the ratings

given in alternate minutes (the minutes were separated by about one week), and from + 73 to + 96, when the items correlated were the means of the ratings given in immediately successive half-minutes

4 The interactivity correlations, even when corrected for attenuation, are not high The corrected coefficients scatter from — 18 to + 71. This fact probably signifies that, in spite of the arbitrariness of the scales, a child's relative standing in activeness is a function, in part at least, of the situation in which he is placed

5 Objective activeness, doubtless, reflects a number of traits, from those with a strong social significance to those whose import is largely physiological.

6. Mental and chronological age seem not strikingly related to any of the measures of activeness achieved, except that of dressing-undressing, and slightly related only to those for pedal (indoors), floor, circle (standing), and apparatus activities.

7 Holding constant chronological and mental age still leaves interactivity correlations of significant size in at least half of the cases.

8 Behavior in the afternoon was less active than in the forenoon The most reasonable interpretation of this fact is not clear.

9 Those lower indoor temperatures obtaining during the observations tended to be associated with less activity than did the higher, whereas the reverse seemed to hold for the outdoor.

10. The humidity out of doors in the range observed apparently *had no effect on activeness.*

11 The boys studied were, as a group, more active than were the girls

References

1 Bridges, K M B. Occupational interests of three-year-old children *J. Genet Psychol.*, 1927, **34**, 415-523

2 Farwell, L Reactions of kindergarten, first- and second-grade children to constructive play materials *Genet Psychol Monog*, 1930, **8**, 430-562

3 Goodenough, F L Inter-relationships in the behavior of young children *Child Develop*, 1930, **1**, 29-48

4 Holzinger, K J Statistical methods for students in education Boston Ginn, 1928 Pp viii+372

5 Sweeny, M E, Hejinian, L. M, & Sholley, R A method for studying the activity of preschool children *J Home Econ*, 1929, **21**, 671-679

University of Chicago
Chicago, Illinois

UNE ÉTUDE DE LA TECHNIQUE DE L'ESTIMATION SURTOUT À L'ÉGARD DE L'ACTIVITÉ DES ENFANTS D'ÂGE PRÉSCOLAIRE

(Résumé)

On a examiné cinquante enfants d'âge préscolaire pendant neuf types d'activités. Environ une fois tous les quatre jours jusqu'à ce qu'on avait obtenu environ trente estimations, on a fait des estimations de l'activité montrée par chaque enfant dans chaque occupation pendant deux demi-minutes successives La corrélation des moyennes des estimations faites les demi-minutes immédiatement successives a donné pour les diverses occupations des coefficients de constance de +0,73 à +0,96, tandis que les coefficients de constance basés sur des minutes d'observation alternées ont varié de +0,44 à +0,74 Il paraît donc que pour un groupe également homogène environ une heure et demie d'observation distribuée correctement en intervalles d'une minute donnerait pour la plupart des occupations des indices d'activité satisfaisants pour un diagnostique individuel On peut montrer que la méthode employée peut donner des résultats pareils quand d'autres expérimentateurs l'emploient, parce que les résultats basés sur les estimations faites par deux observateurs inexpérimentés observant simultanément et indépendemment ont donné des corrélations de +0,90 à +0,97. Les correlations entre les mesures obtenues dans les diverses occupations n'ont pas été élevées cependant (moyenne environ +0,30) même quand on les a corrigées pour l'atténuation—un fait qui suggère le danger d'employer le comportement d'un enfant pendant un type d'occupation comme indice de son comportement pendant tous les autres L'âge mental et l'âge chronologique semblent ne pas avoir beaucoup de relation avec les neuf mesures d'activite achevées, à l'exception d'une. Les garçons étudiés ont été plus actifs que les filles L'activité semble être influencée d'une manière complexe par la temperature mais non pas par l'humidité dans les variations observées

KOCH ET STREIT

EINE UNTERSUCHUNG DES RANGORDNUNGSVERFAHRENS, BESONDERS IN BEZUG AUF DIE TATIGKEIT VORSCHULPFLICHTIGER KINDER

(Referat)

Es wurden 50 vorschulpflichtige Kinder im Verlauf von 9 Arten von Tatigkeiten beobachtet Es wurden ungefahr jede 4 Tage Rangordnungen gemacht (ratings) an der von jedem Kind in jeder Tatigkeit wahrend zwei aufeinanderfolgenden Halbminuten erwiesenen Emsigkeit (activeness), bis ungefahr 30 Urteile erhalten worden waren. Korrelierte man die Durchschnittszahlen der wahrend gleich aufeinanderfolgenden Halbminuten gegebenen Rangordnungen, so erhielt man fur die verschiedenen Tatigkeiten Zuverlässigkeitskoeffizienten von +.73 bis +96, wahrend die auf alternierende Minuten der Beobachtung basierten Zuverlassigkeitskoeffizienten zwischen +44 und +74 schwankten Es scheint also, dass bei einer im gleichen Grad homogenen Gruppe ungefahr 1 1/2 Stunden richtig in einminutige Intervallen verteilter Beobachtung bei den meisten Tatigkeiten Emsigkeitsindexe (indexes of activity) liefern wurden, die zu individueller

Diagnose genugen Dass das verwendete Verfahren in den Handen verschiedener Versuchsleiter ahnliche Resultate liefern kann, wird dadurch erwiesen, dass Zahlen (scores), die auf die Urteile zwei relativ ungeschulter Beobachter, die gleichzeitig und selbständig beobachteten, basiert wurden Korrelationen zwischen +90 und +97 lieferten Die Korrelationen zwischen den an den verschiedenen Tatigkeiten erhaltenen Zahlen waren aber nicht hoch (Durchschnitt ungefahr +30), selbst wenn sie mit Rucksicht auf Attenuirung (attenuation) korrigiert wurden, ein Befund der darauf hinweist, dass es gefahrlich ist, das Benehmen des Kindes bei einer gewissen Tatigkeits als Index seines Benehmens bei allen anderen Tatigkeiten zu gebrauchen Das geistige und das kronologische Alter scheinen mit keiner der entwickelten 9 Massstaben der Emsigkeit mit Ausnahme eines eingelnen, in enger Beziehung zu stehen Die untersuchten Knaben waren emsiger, als die Madchen Die Emsigkeit scheint auf eine komplizierte Weise durch die Temperatur, aber, innerhalb des beobachteten Umfangs, nicht durch die Humiditat beeinflusst zu werden

KOCH UND STREIT

A FURTHER STUDY OF THE INITIAL MAZE BEHAVIOR OF RATS*

From the Psychological Laboratories of the University of Virginia

WAYNE DENNIS, R. H HENNEMAN, AND W. B. VARNER

INTRODUCTION

In an earlier study Dennis and Henneman (4) found that the rat in its initial behavior upon an elevated maze made many fewer errors than would be expected upon a chance basis, and that many of the culs-de-sac were not entered during the entire first run While these facts alone prove the fallacy of the theory that initial behavior is always random and the theory that the rat at first makes all possible errors, they do not determine what factors in the maze situation do control the initial choices. The present paper contributes data concerning the exploration of three additional mazes. These data show that chance, and more than chance, as well as fewer than chance errors may be elicited by a maze In the end we shall attempt to state the factors in the maze pattern which determine initial choices

Not only is this problem legitimate in its own right but it takes on additional interest because of its connection with maze learning. Initial responses are those responses the modification of which constitutes maze learning, yet they have never been adequately studied, in fact, they have even been thrown out of the learning records in order to increase the reliability of the maze!

Probably initial behavior is influenced not only by maze conditions but by many non-maze factors as well, and to describe it thoroughly we ought to measure it as a function of age, hunger, tameness, room temperature, etc. In this study, however, we have chosen to restrict our variation to maze pattern and to hold constant all non-maze conditions with the exception of one slight variation in age

THE MAZES

All of the mazes were of the block elevated type (3) which requires the animal to run along the upturned edge of a wooden block.

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, March 2, 1932.

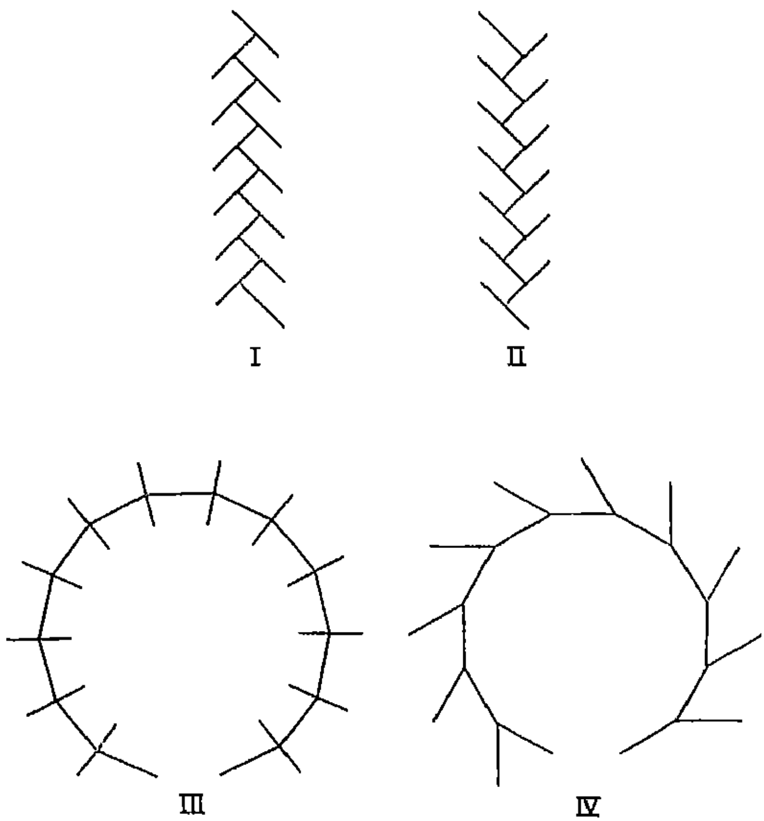


FIGURE 1

The blocks were 2 inches by 12 inches by 24 inches Figure 1 shows the maze patterns. Only Mazes II, III, and IV were used in this study, but Maze I is presented because its data will be compared with the new records which we have obtained.

The blocks were new when used in Maze I Before being used in a different pattern the running surface of each block was thoroughly sandpapered and the blocks were completely "shuffled." The mazes were uniformly illuminated from above, the lamps being placed so as nearly to eliminate shadows. The starting alleys of all mazes occupied approximately the same position in the room and the com-

parative directions of the pathways are indicated by Figure 1 Each maze was set up in the room which contained the rats' cages, the cages being to the upper left of each maze as the reader faces Figure 1, and 10 feet distant from the nearest part of any maze The entrance to Mazes III and IV are to the right of the gaps in the circles.

Animals

Fifty-nine, 43, 54, and 52 albino rats were used on the four mazes, respectively. On Maze I and Maze II only completely untrained rats were used. On Maze III, 13 untrained rats and 41 rats which had previously been placed once upon Maze I were used. Maze IV was explored by 11 untrained rats, 18 rats which had exploratory experience upon Maze III, and 25 rats with exploratory experience upon both I and III This makes a total of 208 explorations by 125 rats, of which 149 explorations and 67 rats are new to this study

All of the rats new to this study were the offspring of the subjects of the previous study, and were born and reared under the same conditions as their parents. They were about 6 months of age at the time of this experiment, whereas the parent rats were 3 months old when they ran Maze I, and 6 months old when they ran the mazes described here

Previous to the experiment all rats were accustomed to have food constantly before them The rats were handled very little. With few exceptions, they had never been outside of their cages. While they ran when one reached for them in the cages, only seldom were they frightened enough to jump from a maze block when placed upon it

Procedure

The rats were given no preliminary feeding nor were they habituated to receive food at a definite time They were, however, deprived of food for 24 hours before experimentation in order to make them more active Each rat was placed upon the first block of the maze and its behavior during the succeeding 7 minutes was recorded by the experimenter who sat behind a screen near the entrance block. The rat was returned to its cage at the end of seven minutes *No food was placed on the maze at any time.* When rats were used on more than one maze at least two days and usually a much longer interval elapsed between trials. All trials were given between 8 and 12 P M While presenting differences, this procedure is essentially the same as that of the previous study

The blocks upon which a rat placed both of its forefeet were recorded in order. The initial "choice" of a pathway therefore means that at a bifurcation the cul or the true path was entered by the forefeet prior to such entry of the other pathway. Through an error in recording technique whether the rat entered the chosen path to a greater degree was not recorded in all cases We can tell whether the *true* sections were entered in full, for if after a recorded choice of the true path the next choice is made at the other end of that block we know that the block must have been traversed, but we cannot tell whether the culs were entered in full. Likewise, the records do not show whether the rat left a blind by merely withdrawing its forefeet or by pivoting and facing a new bifurcation—quite different situations. Consequently, we shall not study the choices made upon emerging from the blinds While this limits our data to choices made from a position in the true pathway, the validity of these data is in no way affected.

Results

1 *Preliminary Considerations* As can be seen in Figure 1, each bifurcation in any maze employed in this study has much in common with the other bifurcations of the same maze. (It will be remembered that we are dealing only with bifurcations met when in the true pathway.) If the initial behavior at any bifurcation is uninfluenced or practically uninfluenced by the behavior at previous bifurcations, we shall be able to simplify our presentation by combining all the data for each maze Table 1 shows that this condition is met. Only in Maze I is there a reliable increase in accuracy in the later sections of the maze, and this increase is only 17 per cent While some error may be introduced into these comparisons by virtue of the fact that many of the rats never reached the second halves of the mazes, they indicate that probably no marked change in accuracy occurs. Data

TABLE 1

Initial Forward Choices in First and Second Halves of Each Maze

Maze	First Half Choices	First Half Percentage Correct	Second Half Choices	Second Half Percentage Correct
I	217	62	159	79
II	133	29	37	32
III	164	57	59	61
IV	162	54	64	47

TABLE 2
INITIAL FORWARD CHOICES

Maze	Choices	Percentage Correct
I	376	69
II	170	30
III	223	58
IV	226	52

are too few for a comparison of the records in first and second halves of those rats which traversed both halves, except in Maze I. As stated in the previous paper, the percentages of accuracy are 65 and 79, number of choices in each half being 159. In the other mazes the choices in the second halves are too few to materially affect the total percentages The early blinds are, of course, those which most affect the total percentages.

2 *Initial Forward Choices.* The total results of the initial forward choices in each maze are shown in Table 2. Initial forward choices represent the behavior of the rat in passing a bifurcation for the first time. Since Mazes I, II, and IV offer at each choice point one true and one false pathway, chance correctness would be 50 per cent On the other hand, Maze III offers two blinds and one true pathway, so that chance accuracy would be 33 1-3 per cent. Table 2 shows that Maze III, as well as Maze I, induces much more than chance correctness (69 per cent and 58 per cent as compared with expectations of 50 per cent and 33 1-3 per cent) Maze II, on the contrary, results in only 30-per-cent accuracy, whereas 50 per cent would be obtained by random behavior. The choices in Maze IV are not reliably different from chance expectation These results indicate that mazes may be so constructed that initial responses will be correct in a proportion either greater than, less than, or equal to, the choices made, say, by flipping a coin

We call attention to the fact that the maze pattern which introduces only 30 per cent correct choices is just the reverse of the pattern which causes the strikingly accurate behavior previously reported The rats are simply introduced into the opposite end of the maze, although the maze is rotated 180 degrees in order to start the animals at the same spot in the room

3. *Interpretation of Initial Forward Choices.* To find what it is in the maze situation which determines the accuracy of the initial

choices is a problem which it is likely that we cannot answer with finality with our limited data Let us begin by eliminating two hypotheses. In the first place, there is plenty of evidence to show that to the untrained rat the visual appearance of a cul has no more inhibiting effect than the true pathway. The chance entrances into the culs of Maze IV and the greater-than-chance entrances into the blinds of Maze II had every opportunity to be inhibited by visual stimuli. To be sure, entrances into blind alleys were inhibited in Mazes I and III, but, in view of the results just mentioned, this cannot be attributed to visual stimulation from the blinds. Likewise, it does not seem that the rat learns to repeat in a new section a turn which has led it forward, for such repetition is possible in all of these mazes, and yet only Maze I shows reliably increased accuracy in later sections.

The orientation of the rat with respect to the various pathways when it faces a bifurcation seems important, as was shown by Dashiell (1) When the rat approaches the choice point in Maze II it is on a block which leads directly into a blind, in Maze III, it is on a block which leads almost straight to a true path This may account for the preponderance of errors and correct choices respectively in these mazes. This also accounts for the fact that in Maze III the inward blinds which require a sharper turn than the outward ones are chosen less often in the initial choices (12 per cent of the time as opposed to 26 per cent) Likewise the equal divergence of the blind and the true paths with respect to the preceding block seems to account for the chance behavior upon Maze IV The rub is that in Maze I the true and the false units are also equally divergent from the direction of the preceding block and yet strongly non-random behavior results

We believe that this exception can be accounted for by introducing a second principle The principle is that a general directional cue may be established which operates beyond the limits of a single block. Where the pathway direction is non-differential with respect to the next true path and cul, this factor is given full sway, as in Maze I. Where the general direction cue is not differential with respect to true path and cul, the direction of the preceding block strongly influences choices, as illustrated in Maze II Evidence for such an initial maintenance of direction has been presented by Dashiell and Bayroff (2) Such an orientation on the initial run might be maintained either by an internal or an external cue, or by both The position of the experimenter's screen with reference to the rat may well introduce a directional element in behavior However, we prefer not to speak of the

general direction stimulus as a forward-going tendency, as Dashiell and Bayroff have done, for the rat *retraces* considerably and our later data will show that the rat is oriented when it retraces. We prefer to think of the rat as guided by some relatively fixed and distant cue, whether proceeding forward or backward. We may think of the experimenter's screen, for instance, as exerting a constant influence, yet the sign of the influence may alternate between attraction and repulsion. The apparent absence of any such directional effects in Maze IV may be due to the complexity of these positive and negative impulses, for at many points in this maze a forward choice would be influenced in the direction of a cul by one impulse, in the direction of the true path by the other impulse, and the relative strength of the impulses may vary from bifurcation to bifurcation. The same is true of Maze III. Unfortunately, the choices at each bifurcation are too few to justify separate treatment. In Maze I the forward impulses apparently overbalance the backward, in Maze II, they are non-differential with respect to forward choices.

We fully recognize the provisional nature of these two hypotheses, but present them as our best interpretation of the present data Of course, behavior may be controlled by other cues, such as experimentally introduced odors (see Watson, 5), but such factors were felt to be absent from this experiment

4 *Initial Retracing Behavior* The initial choices made while moving along the true pathway in the reverse direction agree remarkably well with the forward choices which we have just analyzed Table 3 summarizes the data The choices referred to here are those made when a bifurcation is first reached by retracing in the true pathway.

It has been mentioned earlier that Maze II is the reverse of Maze I. Consequently, when Maze I is retraced, we may expect as high a degree of error as when going forward in Maze II; and vice versa, retracing in II should be as accurate as forward progression in I.

TABLE 3
INITIAL RETRACING CHOICES

Maze	Choices	Percentage Correct
I	55	34
II	97	60
III	144	60
IV	169	63

Table 3 shows that these expectations are reasonably approximated These data substantiate our theory that a reversible directional cue becomes established in Mazes I and II

Maze III, which presents the same alternatives in either direction, produces the same percentage of error in either direction

In Maze IV the retracing situation is novel, but greater-than-chance correctness is to be expected on account of the bodily orientation of the rats at the points of pathway division.

5. *Behavior between Choice Points.* Corroborative evidence of the effect of immediate orientation is found in the behavior of the rats when between two choice points In this situation the rat can continue in its direction or reverse it There is a strong tendency to continue. Of the true blocks which were initially entered in a forward direction, 90 per cent, 84 per cent, 79 per cent, and 84 per cent, respectively, of the units of the four mazes were traversed to their opposite end, leaving only a small number in which the rat turned back Of the units entered while the rat was retracing, 73 per cent, 91 per cent, 94 per cent, and 89 per cent, respectively, were followed to the opposite end, leaving again only a small proportion in which the rat turned and went forward. Since in these mazes the rat had either one or two feet of distance in which to turn, the chances of continuing at any given point are much larger than the figures shown above

6 *Initial Perfect Runs* In the earlier paper it was mentioned that several rats made a large number of consecutive correct turns without retracing, and that each of two rats made as many as 12 such turns The present mazes produced comparable data The most striking performances occurred in Maze III, as might be expected from the high percentage of correctness in this maze One rat made 7, one 9, and one 12 consecutive errorless initial forward choices without intervening retracing In Maze IV there were runs containing 7, 8, 9, 9, and 10 such selections of pathways. In Maze II, on the other hand, the highest number of consecutive correct turns was 5. In our opinion there is a common belief that high accuracy in a maze is always the result of practice These data show this belief to be false, and a little computation will convince the reader that it is highly improbable that the data are due to chance It would be highly interesting to know the sensory control and the permanence of such high initial accuracy Most of our rats have been given only

one trial, but in one case a rat having only one initial error on Maze I was rewarded and ran four additional times on successive days with the result that no subsequent errors appeared

Summary

The initial choices of rats upon three additional mazes have been analyzed. The data show that the initial accuracy of rats in a maze may be greater than, less than, or equal to, random accuracy Culs *per se* are not differentiated from the true pathway, nor does the rat appear to transfer practice in early sections of the maze to later sections to any marked extent It is proposed that two factors may control the rat's initial choice of alternative pathways, namely, the relation of the pathways to the bodily orientations of the rat as controlled by the alley which the rat occupies when the choice is made, and the relation of the pathways to the direction of some more general and more persistent cue This cue may be an external cue such as the experimenter's screen, or it may be intra-organic. Prior evidence of these factors have been presented by Dashiell and by Dashiell and Bayroff The general cue, it is proposed, controls the line of movement but permits the forward or backward directions to alternate In support of this is the fact that in the present mazes retracing choices did not differ markedly from forward choices at comparable bifurcations The effect of bodily orientation is that of maintaining the same direction Attention is again called to the fact that it is possible for initial runs to possess an accuracy that is usually considered to require learning for its establishment

References

1 Dashiell, J E The need for analytical study of the maze problem *Psychobiol*, 1920, **2**, 181-186

2 Dashiell, J F, & Bayroff, A G A forward-going tendency in maze-running. *J Comp Psychol*, 1931, **12**, 77-94

3 Dennis, W A block elevated maze *J Comp Psychol*, 1931, **12**, 429-431

4 Dennis, W, & Henneman, R H The non-random character of initial maze behavior *J Genet Psychol*, 1932, **40**, 396-405

5 Watson, J B Animal education Chicago Univ Chicago Press 1903 Pp 122

University of Virginia
University, Virginia

UNE AUTRE ÉTUDE DU COMPORTEMENT INITIAL DES RATS DANS LE LABYRINTHE

(Résumé)

On a analysé les choix initials des rats dans trois autres labyrinthes. Les données montrent que la précision initiale des rats dans un labyrinthe peut être plus grande ou plus petite qu'une précision faite au hasard ou égale à celle-ci. Les culs-de-sac *per se* ne sont pas distingués du vrai couloir, et il ne paraît pas que le rat transfère l'exercice dans les premières sections du labyrinthe à d'autres sections à un degré marqué. On suggère que deux facteurs gouvernent le choix initial de couloirs alternatifs par le rat, c'est-à-dire, la relation des couloirs à l'orientation du corps du rat gouvernée par le couloir qu'occupe le rat au moment du choix, et la relation des couloirs à la direction de quelque repère plus général et plus persistant. Ce repère peut être un repère extérieur tel que l'écran de l'expérimentateur, ou il peut être intra-organique. Des témoignages antérieurs de ces facteurs ont été présentés par Dashiell et par Dashiell et Bayroff. On suggère que le repère général gouverne la ligne du mouvement mais permet que les directions en avant ou en arrière alternent. Cette suggestion est soutenue par le fait que dans les labyrinthes employés ici les choix de la direction en arrière n'ont pas été très différents de ceux de la direction en avant à des bifurcations comparables. L'effet de l'orientation du corps est celui de maintenir la même direction. On fait rappeler encore une fois qu'il est possible que les parcours initials possèdent une précision pour l'établissement de laquelle l'apprentissage est usuellement considéré nécessaire.

DENNIS, HENNEMAN, ET VARNER

EINE WEITERE UNTERSUCHUNG DER ANFANGLICHEN LABYRINTHTATIGKEIT VON RATTEN

(Referat)

Es sind die anfanglichen Wahlen (choices) von Ratten an drei weiteren Labyrinthen analysiert worden. Die Befunde erweisen, dass die anfangliche Genauigkeit (accuracy) von Ratten in einem Labyrinth grosser oder minder sein kann, als die "blinde" (random) Genauigkeit, oder letzterer gleich sein kann. Sackgassen werden nicht von dem richtigen Pfad unterschieden, und die Ratte scheint auch nicht wesentlich die Übung von den anfanglichen Teilen des Labyrinthes auf die spater kommenden Teile zu ubertragen. Es wird vorgeschlagen, dass die anfangliche Wahl abwechselnder Pfade (alternative pathways) durch die Ratte von zwei Einwirkungen bedingt sein konnte, namentlich, die Beziehung der Pfade zur korperlichen Orientierung der Ratte, unter Kontrollierung dieser Orientierung durch die von der Ratte zur Zeit der Wahl besetzten Allee, und die Beziehung der Pfade zur Richtung irgend einer allgemeineren und bestandigeren Anweisung (cue). Diese Anweisung kann eine ausserliche sein, wie z.B. die Schirmwand des Versuchsleiters, oder sie kann eine innerlich-organische sein. Beweise fur diese Anweisungen sind schon von Dashiell und von Dashiell und Bayroff dargeboten worden. Die allgemeine Anweisung, schlagt man vor, beherrscht die Richtung der Bewegung, gestattet aber die Abwechslung der Richtungen vorwarts oder ruckwarts. Dieser Vorschlag wird unterstutzt durch den Befund, dass in den gegen-

Dieser Vorschlag wird unterstutzt durch den Befund, dass in den gegenwartigen Labyrinthen die Wahlen der Wiederbetretung (retracing choices) sich nicht auffallend von den Wahlen des Vorwartsdringens an vergleichbaren Gabelungen unterschieden Die Einwirkung der korperlichen Orientierung ist die des Fortfahrens in der selben Richtung Man weist aufs Neue darauf hin, dass anfangliche Fahrten (runs) eine Genauigkeit offenbaren konnen, die gewohnlich als nur durch Lernen erzielbar betrachtet wird

DENNIS, HENNEMAN, UND VARNER

THE INVENTIVE GENIUS OF THE CHILD*

From the Psychological Laboratory of the University of Rostock, Germany

ROSA KATZ

This study is concerned with the production of tools and other creative manifestations of our son Julius, born in July, 1922. The production of tools, which subject occupies the first part of this report, reached its high point during the transition from the eighth to the ninth year. The literature of child psychology is everywhere rich in materials regarding the playful work of the child, of its portrayals of objects which are placed before it in the form of available materials such as plasticine, cardboard, paper, wood, building-blocks, and the like. The present study is not concerned with such activities, but rather with genuine tool production, principally out of such materials as stone, wood, and bone which the child himself brings back from a pleasure trip or finds provided in the house, and which he uses in accordance with his present purposes. Stone, wood, and bone are the materials out of which the man of the Stone Age principally made his tools. These materials require and will bear only certain definite preparations, and thus arise, from the hands of the boy, implements and tools which have in part a wholly astounding resemblance to those tools which have survived from the Stone Age. Perhaps the underlying enduring mentality (*Einstellung*), the creative power, of the boy is more interesting than the more or less accidental appearance of the implements. It is just as though the boy, of the age in question, directed to all things in his surroundings, and to all raw materials that come into the sphere of his taste for work, the question, "What can be made out of them? What can one begin with them?" And he undertakes as much as is possible with these materials. There can be no doubt that the pragmatism of the child has lent to the objects of his environment a wholly specific aspect. The original creations arising from the hands of the child are but temporary, and vanish with the moment

*Recommended for publication by David Katz, accepted by Carl Murchison of the Editorial Board, and received in the Editorial Office, February 19, 1932.

I will briefly describe some of my subject's enduring creations in chronological order

In November, 1930, while a goose was being dressed in the kitchen, J. found that the "wind pipe" of the bird could be used as a *water hose,* and caused water to flow through it into a water basin After further handling of this object, he began to blow through it, which produced a hissing sound Thereupon the boy stripped oft the outer tissue, bored a hole through the side of the tube, whereupon the primitive *flute* appeared, not an instrument intended for practical purposes but for play [1] The production of the flute from the wind-pipe was possible only because the wind-pipe had first been used as a water hose, and was consequently moist. J noticed this relationship. From the breast-bone of the goose J. prepared a *mask,* which was secured before the face by means of a cord, with which he usually suspended it from his belt. Neither is the mask a working tool, but an ornament Had the child in this case discovered the primitive mask? The upper thigh-bone of the goose was used for a *club,* so long as the subject was on the war-path In times of peace the club was used as a handle for a cupboard door. This is not the only case observed by us of the altered use of a tool One can say, in general, that the function of a tool is altered in correspondence with a change in the situation

While a joiner was working in the house, J gathered up the wood shavings resulting from the work of repair, sought out the thinnest *ones, moistened them, and bound up a finger, which he had injured* while gathering the shavings Who does not come upon the thought that probably the primitive medicine man used inner bark to bind up wounds, in place of our modern muslin bandages?[2]

After the dismissal of school, J. brought to the house a pocket full of stones. In the house he built a hut of boards, and from the stones constructed a *hearth* in the center of the hut He used the shell of a cocoanut as a cooking utensil From another trip J. brought a paper bag full of modeling clay, out of which he prepared pots His pots were sometimes wholly thick-walled, like the pots of

[1]Pipes are, if not the oldest, in any case very old musical instruments In European, as well as in American graves and caves, one finds pierced bones of slain animals—especially those of birds which do not need to be hollowed out—together with stone tools [C. Stumpf (13, p 242)]

[2]Dr Kurt Haack (Rostock) recently gave me a verbal account of the farmers of *Münsterland* (*Westfalen*) who today still use shavings to bind up injuries

FIGURE 1

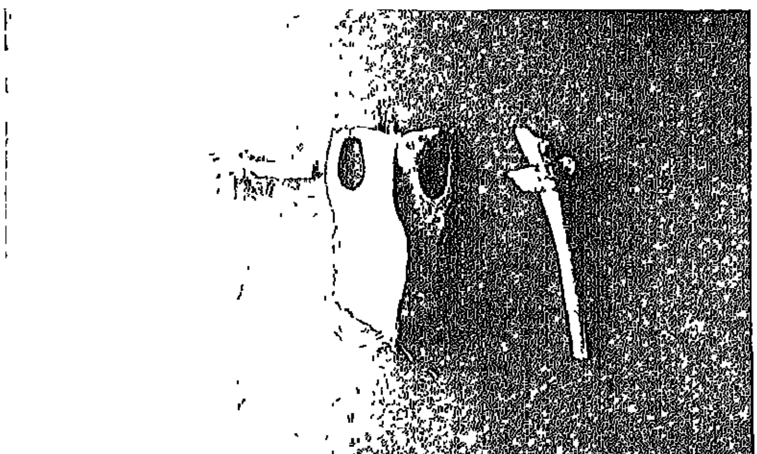


FIGURE 2

old lake-dwellers, while in other cases they were shallow and thin-walled. In any case, he attained great skill as potter. One could have placed these vessels among the pots of a folk-museum without having their want of genuineness observed.

J. split the end of a stick previously made free from bark. He clamped a stone in the cleft and wound the shaft with a cord. There is no doubt that this was the essential appearance of the first axe swung by the human hand. According to Pfeifer (9, p. 186, 10, p. 99), this is just the way American and Australian shafts were prepared. *J. made a milling stone out of a large flintstone naturally* possessing the hollow form required.[1] With this crucible J. ground up the kernels of hazelnuts to the consistency of flour, which he ate with great relish. He used a flintstone with a sharp edge as a scraper,[3] using it to remove bark from branches. The subject impaled a short stick with a hole through its center upon another long stick with a sharpened end, the whole being festively decorated with a cord. It was a tomahawk, which, during peaceful times, served as a little hammer. When J. acquired a piece of bacon containing a flattened bone about 2 centimeters in breadth, he cut out the bone and used it in its natural form as a spoon (*Schieber*) while eating half-fluid foods. It was entirely suited to this function. A throwing weapon or *spear* was produced by means of a sharpened stick with a cork impaled on it. J. wanted a girdle, and, finding two short straps insufficient when taken alone, bound them together in a skillful way by means of two binding pieces taken from a box of metal building materials. Petals were placed in an empty match box: this served J. as a little smelling vessel. The marrow was eaten from a hollow bone about four to five cm. in length by means of a suitable bone spoon especially sought out for this purpose. J. used a sewing thread to which chicken feathers were secured for a garland. An empty snail-shell, strung upon a feather which secured it to the thread, served as the center-piece. The snail-shell was used as a pendant during the Stone Age.

In July, 1931, while at the ocean beach, J., burying a bottle of cocoa in the sand, put a straw through a hole in the paper cover, bent over the bottle and drew out the cocoa. This was drinking out of a spring or out of a brook, as primitive man did it. No other child

[3]Pfeifer: "Immediate and personal use of an object is significant. Here belong *Eolithe* whose development into a typical form of knife and scraping tool, often repeated, is described in an extensive literature" (10, p. 2).

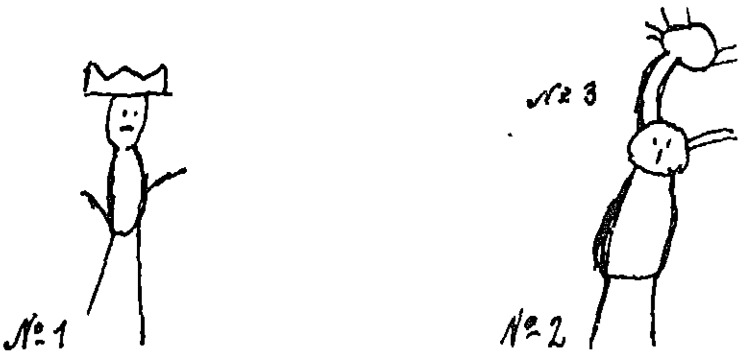
upon the beach hit upon the idea of thus burying and emptying a bottle. In October, 1931, a seven-year-old American child, O W, who understood no German, came to Rostock While playing with J, who understands no English, O. wound up the former's dearest toy, a locomotive, too rapidly. J took the toy, and said, *"Pass auf, jetzt mache ich"* Whereupon J. wound the locomotive carefully, kissed it tenderly, meanwhile never taking his eyes from O. to make sure that his American friend would see the gesture of his love, of his affection for this locomotive

During another time when J was playing with O, the following drama was enacted

J Wie heist Diener, Knecht? (no one answers) Oliver, Du bist niedrig (signifying with his hand, low, little), ich bin Majestat Ich bin kin (he wants to say, king), you ist magd

O What do you mean?

J Wo ist Papier? (paper and pencil are provided) Ich male es, Oliver, komm. That bin ich, Konig (points with the index finger at himself and then at drawing Number 1), that Oliver, Diener (points toward O and then at drawing Number 2), weisst du, was es bedeutet? Ich drucke ihn unter (points to drawing Number 3) Versta?



Like a primitive person, J here seizes upon the sign of suppression (*Drucken*) which is to document the relation of a king to his servant

The foregoing may suffice as a representation of the general genius of discovery in the child and of the particular sort of tools described There remains the task of making the experience of the child understandable in an individual manner, and to carry through some theoret-

ical reflections regarding the childish production of tools in its relation to the production and use of tools by primitive man, by *homo faber*.

Are we, indeed, when dealing with the creations of this child, dealing with original capacities; have not the models of the museum as well as pictures which came under the subject's regard played a rôle, and have not portrayals previously found in reading had their initial influence? It is not easy to determine whether and how much this occurs. For certain cases this explanation decidedly does not apply: for the flute, the mask, the spoon, the smelling box, the necklace, the bandage. Visits to museums with ethnological divisions lie years in the past. J. has, however, had opportunity to learn about the use of primitive tools from reading, of which the most important source was *Robinson Crusoe*, which the children have read, in various editions and different pictures for years. I am inclined to think that this has acted less as a particular stimulus to produce the tools just described than it has acted as a general stimulus to make something out of available materials with which one can satisfy purposes, to produce something with which one can help oneself in difficult situations. Innumerable children read *Robinson Crusoe*, go to museums, and could have acquired such stimulation, but few have the talent, which must be present, for a production to take place.[4] And, therefore, I believe that the individual side of the case will be clarified by reference to the fact that J. is a personality of self-maintained and outspoken practical gifts. For this assumption our day-books supply many evidences, showing that the activities of the child acquire proper illumination from earlier times, preceding the production of the tools, where many individual tendencies are seen to interlock, giving a unified personal picture.[5] A study of the behavior of the child during the telling of mythical stories (6) shows that J. belongs to that class of children who do not want always to hear the same story again. He wants to experience something new and not always hear the same tale. From studies of touch it appears that J. belongs to the group having the shortest reaction-time (8). In a study of the invectives of children (5) J. made the most original contributions. He discovered invectives which he did not and could

[4]"Only he who carries the technical, scientific, or art problem in himself, can seize upon a chance event and see its worth, which he utilizes in finding a solution" [Selz (11, p. 238)]. Stern expresses it in the sentence, "Every tendency to play is the dawn of a serious instinct" (12, p. 265).

[5]Compare, D. and R. Katz (3, 2).

never have heard in his surroundings For example, at the age of four, he coined, while under strong emotion, the invective, "You dirty towel!" which he directed at the children's governess. This is an innovation because this term does not occur in the vocabulary of adults as a term of disapprobation

J. also belongs to the class of children who test their powers as poets, and, though they are few in J.'s case, still such attempts win a certain significance in view of the foregoing [6]

Following is an example of J.'s capacity as poet at the age of four years

> Es war einmal ein Mann
> Da ging er an den Strand,
> Und ass dort Brot und Sand

At the age of eight, J composed "The Flower-Girl"

> Was duftet's hier im Sonnenschein
> Hier sind wohl lauter Blumen fein
> Die Fuchsien schon und gut gebaut
> Die werden von ihr nicht gehaut
> Am Schreibtisch sitzend sie betracht
> Die Blumen und die schone Pracht

J also invented stories of the following type:

> "A hunter went to the forest. One day an eagle came The eagle perched upon a crocodile that was in the water The hunter was frightened when he saw the crocodile and collapsed into the water The crocodile was going to swallow him The hunter seized the eagle's legs, the eagle flew up, carrying the man along, higher and higher, until earth was reached"[7]

Though this production does not stand at a high artistic level, it is nevertheless a personal invention [8] Before going to sleep and before rising in the morning, J fabricates tales which apparently consist of eidetic images Upon my special request, he shared with me the following phantasy:

---

[6]Compare herewith Karl Groos: "Approximately at the age of four one can often observe that children make the attempt to express themselves in the form of poetic style, thus appearing as producing artists" (1, pp 41-42) See also K Tschukowski (14), who presents a large number of verses of four-year-olds.

[7]W Stern "Lively children are not satisfied by merely hearing stories, they want to produce them themselves, and so begin to fabricate" (12, p 296).

[8]An example of the *"fremdbezugliche"* type of fable, as Stern calls it.

"I dream that the children living below us in the house are our sisters, and the others living in the neighboring house are also our sisters Father has written it on a piece of paper that they are our sisters; and that Theodor (J.'s brother) and I are Indians and have much gold and silver The story never comes to an end, I always think of something in addition, it continually goes further."[9]

The *inclination to the artistically perceptual also shows itself* in the rich production of pictures which begins with the close of the fifth year J. produced many tasteful black and white drawings, colored pencil drawings, as well as water colors Some of these productions have a story-like character, others are of ornamental nature At this time, J. was at war with orthography, since he inclined to the creative in this respect also.[10]

Already while in the kindergarten and later J often returned from school with considerable tardiness, because, as he said, "other ways" and "new ways" home were sought. With this, again in obedience to the impulse not to be satisfied with the traditional, and to make new discoveries, J. showed himself inventive in the preparation of new dishes He was not satisfied with what was set before him, but combined foods in a wholly remarkable way, which preparations he ate with the greatest relish.[11]

J is inexhaustible in the discovery of new forms of clothing. Dressing, however, often resolves itself into merely undressing, when he wishes to appear as a primitive man, or as one can say with reference to his inner attitude, in which by preference he *is* primitive man. In general, J. has a great capacity for illusions, as, for example, while at the sea coast in 1926 he made cakes out of sand, putting them into his mouth to eat them [12]

[9]This belongs to the class of "*dauernde Einstellung*," as Stern (12, p 296) calls them A similar case of a long-spun tale is the experience of the youth, Goethe, who, as he relates in his autobiography, always had to tell continuations of stories to his companions J. however does not tell his stories willingly

[10]It is to be observed that educational requirements were lightened as much as possible for J during the first three years, in consequence of which his individual life could develop and be extended. "Before the compulsions of education begin the lives of children consist, excepting the needs of nourishment and sleep, almost entirely of playful occupations Play presents itself to us here as a unified, all-sufficing force of life, indeed as the single life-motive of the child" Groos (1, p. 478)

[11]Completer details in D and R. Katz (4, p 313)

[12]Compare herewith R Katz (7, p 495), and Groos (1, p 394) where entirely the same instance is described

I may believe that the foregoing data are sufficient to support my assumption that J is a child personality that inclines to creative and inventive originality. His impulse to deeds is satisfied in harmony with the prevalence and development of temporarily ruling functions, sometimes in this field, sometimes in that. But always something personal appears, although naturally on a wholly childish level[13] His strength therewith lay, not upon the logical theoretical side, but wholly on the side of the practical-intuitive The highest which he so far has accomplished is the production of tools What raises this individual case above casuistry, and what makes his case worth reporting is the suggested possibility of expanding these parallelisms with the facts of ethnology to a general principle.

With this I come upon the last point of this discussion, namely, the relation of childish tool invention to the invention of tools by *homo faber* The inventions of our boy do not stand in accidental and disconnected association, but show the closest interdependence since they are created from his particular creative gifts The child, as well as the adult, of a cultured folk need have no cares regarding the satisfactions of individual needs. The rule is that the means for satisfying the needs of the person of culture are offered before these needs themselves arise. Through his inclination to the primitive, and through his particular talents, J found the means of personally satisfying his fictive needs I wish then to suggest the thesis that the man of nature, under hard necessity of finding satisfaction for his needs, derived them in a way similar to that of J —out of his own capacities It could well be a general result of ethnological research that becoming human consists essentially in the invention and conscious use of tools. Dawning intelligence was practical intelligence and occupied itself with the satisfaction of needs by introducing things of the surrounding world into personal use The need for nourishment rises above the animal level with the first refinements in securing tools and provisions of plant and animal nourishment Threats of welfare and life through animal and human enemies led to the appearance of weapons of offense and defense Protection against weather evolved clothing and building technique Primitive man still has this in common with the animal, that he tests the

[13]Groos " It can hardly be doubted that these trials of various possibilities often can have the most decisive effect on the later choice of a calling, in that play initially starts the unfolding of the innermost talents and tendencies of character" (1, p. 393).

edibility of all things in his environment, and has exercised this activity tirelessly. Just this practical inclination permitted the things of his environment to be tested regarding their use in satisfying his other needs The naturally given materials were stones, wood, the bones, as well as the other unedible and unseizable parts of dissected animals. Sometimes a tool would appear in the phantasy of a primitive person before suitable material was found which made its actualization possible. But just as often, or perhaps oftener, the case would arise, that only after the accidental encounter with such and such shaped material, the thought of its usefulness would appear J did not seek a bandage, but the wood shavings became suitable in consequence of the need arising from being wounded. J did not seek the mask, but the breast bone of the goose looked like a face, and so it was easy to recognize its mask-function Just as J, because of his disposition, was constantly on the alert,[14] so was it also with primitive man with his necessarily practical attitude And just as playthings can be created because of J.'s capacities, so we can be certain of this, that the primitive person on the search for objects of a practical worth created toys quite early, upon which aimless fancy could exercise itself Finally, as already mentioned, as with J's use of his creations, with primitive man also the same tools underwent changes of function and of use with the change from one situation to another

References

1 Groos, K Die Spiele der Menschen. Jena Fischer, 1899 Pp viii+538

2 Katz, D, & Katz, R. Die Erziehung im vorschulpflichtigen Alter Leipzig Quelle u Meyer, 1925 Pp 134

3 ———. Gespräche mit Kindern Berlin Springer, 1928 Pp vi+299

4 ——— Psychologische Untersuchungen über Hunger und Appetit *Arch. f d ges Psychol*, 1928, 65, 292-320

5 ———. Schimpfworte des Kindes *Beihefte z Zsch. f angew. Psychol*, 1931, 17, 120-131

6 Katz, R. Beobachtungen an Kindern beim Märchenerzählen *Zsch f angew Psychol*, 1926, 28, 140-143

7 ——— Neue Formen des Kindergartens *Zsch f päd. Psychol*, 1927, 28, 491-509

[14]Pfeifer (9, p. 117) has arrived at the conclusion that the hunter of the Stone Age used the same instrument now to saw, now to cut, now to scrape, now to whet, according to need He had at all events executed with one instrument tasks for which we would regard two or three special tools as necessary.

8 ——— Das Tasten des Kindes. *Arch. f. d ges Psychol*, 1930, 77, 35-49

9 PFEIFER, — Die steinzeitliche Technik Jena Fischer, 1912.

10 ——— Die Werkzeuge des Steinzeitmenschen. Jena. Fischer, 1920

11 SEITZ, O Der Schopferische Mensch. *Zsch f Psychol*, 1903, 32, 238

12 STERN, W Psychologie der fruhen Kindheit (5 Aufl) Leipzig Quelle u Meyer, 1928 Pp xii+539

13 STUMPF, C Philosophische Reden und Vortrage Leipzig Barth, 1910 Pp 261

14 TSCHUKOWSKI, K Malenkije Deti Leningrad, 1928

University of Rostock
Rostock, Germany

LE GÉNIE INVENTIF DE L'ENFANT

(Résumé)

Dans cette étude il s'agit de la production des outils et d'autres manifestations créatrices du fils de l'auteur, lequel est décrit comme du type pratique-intuitif plutôt que du type logique-théorique L'importance du cas est la possibilité suggérée de developper les parallélismes entre son developpement et les faits de l'ethnologie jusqu'à un principe genéral Par exemple, l'enfant a utilisé pour ses besoins ce qu'il a trouve dans la nature, ou s'est servi des choses comme il a vu qu'elles pourraient être utilisees, de la même manière dont l'homme primitif a satisfait ses besoins Les emplois de "paix" et de "guerre" des mêmes instruments ont différé avec l'enfant comme on les rapporte au cas de l'homme primitif

KATZ

DAS ERFINDUNGSTALENT DES KINDES

(Referat)

Diese Untersuchung besichtigt die Schaffung von Werkzeugen und andere erfinderische Tätigkeiten des Sohnes des Verfassers, der als ein Kind von praktisch-intuitiver eher als von logisch-theoretischer Art beschrieben wird Die Wichtigkeit des Falls liegt in der suggerierten Moglichkeit der Erweiterung der Parallele zwischen seiner Entwicklung und den Tatsachen der Ethnologie in ein Allgemeinprinzip. Das Kind wendete zum Beispiel das, was er in der Natur fand, zu seinen Zwecken an, in der selben Weise, auf die der primitive Mensch notwendige weise seine Bedurfnisse befriedigte Die "friedlichten" und die "kriegerischen" Nutzen der selben Werkzeuge waren bei dem Kind verschieden, wie sie es, den Berichten nach, beim primitiven Menschen sind

KATZ

THE IMAGINARY PLAYMATE: A QUESTIONNAIRE STUDY*

From the Psychological Laboratories of Columbia University

E. B. HURLOCK AND M BURSTEIN†

The phenomenon of the imaginary playmate occupies one of the vast, unexplored fields of child psychology. The reason for this apparent neglect is obvious to anyone who studies the problem The imaginary playmate is created by the child exclusively for his enjoyment. In this world of make-believe, the older, more sophisticated "grown-up" is regarded as an intruder. For this reason, children are reticent about discussing their imaginary fancies with their elders, and even parents are only occasionally allowed a glimpse of this precious world.

For a psychologist to get first-hand information from the child himself is, therefore, exceedingly difficult. The only alternative of this method has been to obtain reminiscences of adults who, as children, experienced imaginary comrades. Because of the drastic effects of time on memory, this method is subject to a serious error Therefore, most of the work previously reported along this line consists of cases familiar to the psychologist with little or no attempt to reduce the data to an objective form.

Harvey (7) is one of the most recent writers on the subject of the imaginary playmate, and it is he who is largely responsible for having brought the matter to the front pages of child psychology His material consists primarily of a number of case studies, the histories of which he knew personally The nearest approach to a systematic procedure is found in the study made by Vostrovsky (14) in 1890 The material used by her was derived from (1) 27 persons who had imaginary playmates, (2) 10 people who had observed children playing with these created beings, and, (3) 5 people who had heard about such children

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, March 18, 1932

†The data reported in this study was obtained by the junior author for the Master of Arts essay The senior author directed the study, aided in the interpretation of results and in the presentation of the material

While it is true that many other psychologists have dealt with this problem in their writings, none have given the careful attention to the matter that the two above-mentioned writers have done. Certain factors stand out in connection with this study about which all of them agree. On the other hand, there are essentials that show great divergence of opinion. For this reason, it seems advisable to present the facts in a topical rather than chronological sequence.

1. *Reality of the Imaginary Playmate to the Child.* In all of the cases described by him, Green (5, p. 11) depicts the reality of the imaginary playmate to the child.

Harvey (7) states that children have playmates "which are wholly imaginary, but which are as vivid and real to them as living playmates would be. These playmates are not merely vivid ideas, or imaginings, but actual visual and auditory projections. They can be seen and heard as vividly as if they were living children."

Martin (9), Munroe (10), Norsworthy and Whitley (11), and Vostrovsky (14) all stress, in their writings, the vividness and reality of the imaginary playmate to the child.

2. *Personality of the Playmate.* In most of the cases cited by Green (5), the imaginary companions have distinct appearances and definite names applied to them. According to Martin (9), the personalities of these illusory beings are so definite that the child later identifies real people with them. Kirkpatrick (8) believes that "sometimes the imaginary companion is an ideal self, sometimes a naughty scapegrace, and, at other times, not self at all, but a distinct personality. The same child may have many such companions at once, or one at a time in succession."

Vostrovsky (14), in speaking of these companions, says: "Many are great, beautiful and rich, some peculiar, as the companion of a little boy 'that was about three feet tall' and who seemed to him 'as if he could not speak nor hear', others are helpful and kind to whom all troubles are confided; while still others are to be helped in some way." In regard to names, Vostrovsky (14) points out that most children attach specific names to these companions. The names are either common-place, everyday ones, such as "Jack," "Alice Davids," or strange unusual ones invented by the child. The former occur almost three times as often as the latter.

3. *Age of Child at First Appearance of Playmate.* Vostrovsky (14) found that the first appearance of the imaginary playmate varied from the first to the thirteenth year, while Smith (12), Harvey (7),

Martin (9) and Norsworthy and Whitley (11, p 163) all stress the fact that these companions appear very early, usually at the time when the child is beginning to remember things. Munroe (10) and Kirkpatrick (8) give the specific ages of three and four years, the age when the dramatic tendency usually begins. The creation of the imaginary companion seems to them to be an outgrowth of the general tendency to dramatize which reaches its climax at about the age of seven.

4. *Age of Child at Last Appearance of the Playmate* Smith (12), Tanner (13), and Arlitt (1) stress the fact that the imaginary companion usually disappears when the child goes to school and becomes absorbed in outside things or other playmates Harvey (7) says that "an examination of the accounts of different persons who have experienced imaginary playmates indicates that most of them disappear principally at two periods There is a tendency in the reports to fix the age at which they disappear either at the age of seven or eight, or at the age of eleven or twelve." Vostrovsky (14) found that the age for the last appearance of the imaginary companion varied, in her subjects, from seven to seventeen years of age

5 *Type of Child Who Experiences the Imaginary Companion.* Brittain (2) states that "on the whole those who have had imaginary companions have been above the average in the imaginative quality of their stories" Burnham (3) believes the same to be true, while Harvey (7) goes so far as to state that "no stupid child ever had an imaginary companion." Kirkpatrick believes that "not only do a few lonely and highly imaginative children have these companions, but nearly all children have them in some form for a greater or lesser period of time" (8, p 168), and Smith (12) and Tanner (13, p 206) believe that it is the lonely and imaginative child who creates phantom companions as his playmates.

6. *Explanation of the Phenomenon.* In trying to explain the existence of the imaginary companion, many theories have been given Arlitt (1, p 159) thinks that the imaginary companion is created for the child's enjoyment, while Burnham (3, p 212) goes so far as to state that imaginative children not only have imaginary companions but they also personate animals and men. Green (5) includes this phenomenon in the category of day-dreams Harvey (7, p. 1) states that "an imaginary playmate is a visual or auditory idea that becomes as vivid or real as a visual or auditory percept would be . The child recognizes that they differ from real,

living, objective children." Smith (12) suggests that "closely akin to the story form of day-dreams is the imaginary conversation which is sometimes carried on with actual friends and acquaintances, sometimes with strangers casually seen, or with children in history or books, or in some cases with purely imaginary characters."

Procedure

The procedure used in the study to be reported in the following pages was simple and direct. A questionnaire,[1] aimed at three classifications, (1) the commonness of the phenomenon of imaginary companions, (2) the background of the child, and (3) facts about the imaginary playmate, was answered by 701 high-school and college students. Of this number, 393 were women, and 308, men. The ages varied from 15 to 40 years, with the median at a point between 18 and 19 years.

Results

1. *Commonness of the Occurrence of the Phenomenon.* Of the total number of subjects questioned, 31% of the women and 23% of the men reported that they distinctly remembered having had an imaginary playmate. Among the women, 50% testified that they maintained a constant companionship with this companion, 46%, that they played with the imaginary comrade more than once; and only 4% remembered having seen this apparition but once. In the case of the men, 53% played with this phantom friend more than once, and 46% maintained a steady friendship with him. Thirty-eight per cent of the girls and 20% of the boys had more than one imaginary comrade.

2. *Comparison of the Backgrounds of the Subjects.* The background of the child who has had imaginary companions was then compared with that of the child who has not experienced this phenomenon. The results are presented in Table 1.

An examination of this table yields rather negative results. Apparently the size of the family does not influence the creation of imaginary companions. There are as many small families among those who have not had imaginary companions as among those who have had them. Only children were not found to be more subject to imaginary companions than were those who had brothers and sisters.

[1]A copy of the questionnaire is printed at the end of this paper.

TABLE 1
COMPARISON OF THE BACKGROUND OF THE SUBJECTS

	Females who have had the imaginary companion	Females who have not had the imaginary companion	Males who have had the imaginary companion	Males who have not had the imaginary companion
1. Number of children in family				
Subject is only child	21%	13%	9%	10%
One more in family	20%	28%	28%	24%
Two more	22%	19%	24%	19%
Three more	14%	15%	16%	15%
Four more	7%	12%	9%	10%
More than four	16%	13%	14%	22%
2. Place where childhood was spent				
City	85%	83%	83%	88%
Country	15%	17%	17%	12%
3 Parental care				
Both parents	88%	92%	93%	95%
Mother only	9%	5%	3%	3%
Father only	1%	1%	—	1%
Relatives	2%	2%	3%	1%
Institution	—	—	1%	—
4 Number of comrades in childhood				
1 - 3	27%	15%	16%	13%
4 - 7	39%	46%	20%	31%
8 - 10	16%	21%	34%	30%
More than 10	18%	18%	30%	26%
5. Interests and leisure time activities				
Reading	44%	41%	37%	42%
Athletics	24%	28%	32%	37%
Music and art	12%	19%	14%	9%
Miscellaneous	20%	12%	17%	12%

Although among the women there is a slight difference in favor of only children in the group that has experienced this phenomenon, the difference is refuted in the masculine group where a small difference is found in favor of those who have never had an imaginary comrade.

The locality in which one has spent childhood days does not seem

to bear any relationship to this phenomenon. Among the men, 17% of those who had experienced imaginary companions came from rural districts and only 12% of those who did not have the companions In the case of the women, 17% of those who did not have imaginary companions lived in the country as compared with 15% of those who had this unreal companionship

3 *Facts Concerning the Imaginary Playmate* According to data obtained in this study, the imaginary companion does not appear as early in the life of the child as has been supposed by many psychologists. Among the girls, the age at which the imaginary companion is most likely to appear is between 5 and 7 years of age Boys experienced this phenomenon at a considerably later age than did girls One-third of the group fixed the age of the first appearance of the imaginary playmate at the stage between 7 and 9 years of age.

TABLE 2

OCCASION FOR THE FIRST APPEARANCE OF THE IMAGINARY COMRADE

Occasion	Females	Males
Companion just came suddenly	42%	45%
Being lonesome	21%	20%
Having been scolded	3%	3%
Not being able to get along with real friends	3%	9%
Character from story-book or picture	20%	9%
Character from dream	8%	7%
Being unjustly treated	3%	7%

Table 2 shows the suggested reason for the first appearance of the imaginary playmate. Forty-two per cent of the girls and 45% of the boys testified that the apparition "just came suddenly" with no adequate cause for it On the other hand, 21% of the girls and 20% of the boys stated that "being lonesome" caused the first appearance of the imaginary playmate In the case of 20% of the girls and 9% of the boys, the origin of the imaginary companion was traced to some story-book character or picture

Table 3 shows the range of emotions aroused by the playmate in his creator About 80% of the emotions aroused in both boys and girls were the kindly ones, such as love, respect, sympathy, kindliness,

TABLE 3
EMOTIONS AROUSED BY THE IMAGINARY PLAYMATE

Emotion	Females	Males
Love	45%	34%
Hatred	—	—
Fear	1%	—
Respect	12%	20%
Rivalry	3%	11%
Anger	—	—
Mastery	3%	4%
Submission	2%	1%
Kindliness	12%	9%
Contempt	2%	—
Sympathy	12%	11%
Curiosity	6%	7%
Companionship	2%	3%

and companionship There is not one instance of either hatred or anger existing between the creator and his puppet.

The child's attachment to the imaginary playmate is so great that 38% of the girls and 42% of the boys actually preferred these phantom playmates to any real companions The child loves his imaginary playmate and is apt to attach great secrecy to him Among the girls, only 21% were willing to share these companions with anyone. Much greater diffidence appears to exist among the boys. Only 17% of the boys were willing to share these companions, and but 6% discussed their doings with other people.

The imaginary companion is a distinct personality to the child who experiences it. As many as 81% of the girls and 60% of the boys testify that the playmate was real to them. Seventy-nine per cent of the former and 43% of the latter conversed with this comrade "imagining they heard him talk." In studying the names attached to these companions, it was found that the common ones, Ruth, Pauline, Sarah, Robert, Oswald, etc , predominated In the case of the girls, 83% of them, as contrasted with 69% of the boys, attached definite names to these companions.

Table 4 shows the reasons for the choice of a name for the imaginary companion. Thirty per cent of the girls and 20% of the boys bestowed upon the imaginary companions the names they liked best in real life. The girls were found to choose names from story-book characters in 33% of the cases, while among the boys the influence of the cinema was greater.

TABLE 4
REASON FOR CHOICE OF THE NAME OF THE IMAGINARY PLAYMATE

Reason	Females	Males
No reason	33%	55%
Name of someone liked	30%	20%
Name of someone disliked	——	——
Name of hero or heroine worshipped	2%	8%
Name of story-book character or picture	33%	5%
Name of a movie star	2%	12%

Children were found not to crave companions of essentially the same ages as they Their imaginary companion, like the real companion, is rarely ever younger than his creator There is a definite tendency among both boys and girls to create companions who are older than they Much the same was found to be true in the case of the sex of the imaginary companions. More than half of the subjects had real companions of their own sex and created imaginary companions who were also of the same sex. Not one boy and only 8% of the girls who enjoyed the real companionship of the opposite sex cared to create an imaginary companion of the same sex

The prevailing view of the psychologists cited above seems to be that the child loses his imaginary companion when he enters school, or at least, as he matures. Forty-eight per cent of the boys were found to be carrying on the imaginary friendship after the age of ten,

TABLE 5
OCCASION OF THE LAST APPEARANCE OF THE IMAGINARY COMRADE

Occasion	Females	Males
Entrance into school	11%	20%
Companionship of other children	33%	33%
Sudden disappearance	44%	13%
Interference of outsiders	3%	7%
Disagreement with imaginary companion	1%	——
Shamefulness	3%	——
Entrance into college	1%	——
Falling in love with real playmates	3%	——
Imaginary companion "died"	1%	——

while one-fourth of the girls brought the imaginary companion with them into adolescence. In most of the subjects, the occasion for the last appearance of the imaginary playmate was hazy. In Table 5 are presented reasons for the last appearance of the imaginary comrade.

Summary and Conclusions

1. The creation of the imaginary companion seems to occur more often among girls than among boys. More than half the girls and almost as many boys have maintained a constant companionship with their imaginary friends. Almost two-fifths of the girls and one-fifth of the boys have had more than one imaginary companion.

2. According to the indices used in this investigation, the background of the child who has experienced the imaginary comrade does not differ materially from that of the one who has not had these friends. The information secured points to the need for more extensive investigation in this field.

3. The age at which the imaginary companion makes its first appearance is considerably older than has been supposed. Among the girls, the most frequent time is the age between five and seven. A great number experience the imaginary playmate for the first time after the age of ten. The boys experience the phenomenon at a considerably later period; the most frequent stage, here, is after the age of ten.

4. The occasion for the first appearance is indefinite in the minds of almost half the subjects. Where the reasons are stated, "lonesomeness," and inspiration from a "story-book character" are the causes most frequently ascribed.

5. In general, the companion is very much treasured by the child. In four-fifths of the cases, the companion aroused kindly emotions in the creator; there is no instance of either hatred or anger existing between the creator and his creation.

6. In general, the child does not like to share his imaginary companions, nor does he care to discuss their doings with other people. Boys show greater reticence than girls in this matter.

7. The imaginary companion is a distinct personality to the child who creates him. In most cases, this comrade can be seen and heard as if he were real. Girls seem to be more under the influence of the illusion than boys.

8. Girls are, also, more apt to attach names to their imaginary

companions than boys are. The reason for the choice of a name is indefinite in most of the cases, but more so in the case of the boys than with the girls The reason most frequently given is the name of someone liked, or of a story-book character

9 It appears that there is no definite tendency in a child to create children of either his own age or sex, although there is evidence that the child does not like to have either living or imaginary companions who are younger than he.

10. The last appearance of the imaginary comrade occurs much later among the boys than among the girls Among both groups the most frequent time is after ten years of age One-fourth of the girls and almost half of the boys have maintained this friendship up to the present time

11 Very few of the subjects can assign definite reasons for the last appearance of the imaginary companion, but this is especially true of the male group Where reasons are given, the most significant one seems to be the companionship of other children.

Questionnaire

Columbia University, New York City

Please answer the following questions as truthfully and as accurately as you can In most cases, you have to check the correct word, in a few others, one-word answers are required. You need not put your name on the paper, as we are interested in group records rather than individual cases

Sex Age , , Grade in School
Subject of major interest in School
Hobby or special leisure-time activity

1 Are you an only child? Yes No
If not, specify the number of younger brothers.
older brothers
younger sisters
older sisters

2 Where did you spend your childhood? City Country

3 With whom did you live during childhood? Both parents
Father only
Mother only
Relatives
Strangers
Institution

4 How old were you when you entered the elementary school?

5 As a child did you have many playmates? Yes No

6 Approximately how many companions did you have?

7 In general, were your companions older, younger, or of the same age as yourself?
Older Younger , Same age

8 Were most of your childhood friends girls or boys?
Girls Boys

9 Did you ever have an imaginary playmate? Yes No

1. How old were you when the imaginary playmate appeared to you for the first time?
2 3 4 5 6 7 8 9 10 Over 10
2 What was the occasion for the first appearance?
Just came naturally
Being lonesome
Having been scolded
Being unjustly treated
Not being able to get along with real friends
Character from story or picture
Character from a dream
3 Did this imaginary comrade appear to you once, more than once, or as steady companion?
Once More than once Steady companion
4 What were the occasions, if any, of the subsequent appearances?
Same as for first appearance
Desire to see companion again
Companion just appeared suddenly
5 Was this playmate real to you? Yes No
6 Did you converse with this comrade imagining that you heard him talk? Yes No
7 Did you see this playmate as clearly as if he were real?
Yes No
8 What was the sex of the companion? Girl Boy
9 Was this playmate older, younger, or of the same age as you?
Older Younger Same age
10 What was the name of this playmate?
11 Can you give a reason for this name?
No reason
Name of someone liked
Name of someone disliked
Name of hero or heroine worshipped
Name of story-book character
Name of a movie star
12 What emotion did this playmate arouse in you?
Love Hatred Fear Respect Rivalry
Anger Mastery Submission Kindliness
Contempt Sympathy Curiosity
13 Did you prefer to play wth this imaginary comrade rather than with real playmates? Yes No
Or did this imaginary playmate come to you only when you could not have the company of real children? Yes No
14 Does the playmate ever appear to you now? Yes No
If not, approximately how old were you when this playmate ceased to appear?
5 6 7 8 9 10 Over 10
15 Did any of your real companions share this imaginary playmate?
Yes No
16 Did you like to speak to older people about this imaginary playmate? Yes No

17 Did you have more than one imaginary playmate?
Yes No

18 In what respects did this second playmate resemble the first one?
Age Sex Appearance Emotions aroused

19 What was the occasion of the last appearance of the imaginary playmate or mates?
Entrance into school
Companionship of other children
Sudden disappearance
Interference of outsiders
Disagreement with the imaginary playmate
Shamefulness

20 Any additional incidents you would like to list in connection with the imaginary playmate not covered in the above

References

1. ARLITT, A. H Psychology of infancy and early childhood New York McGraw-Hill, 1928, Pp xi+228

2 BRITTAIN, H L A study in imagination *Ped. Sem*, 1907, **14**, 137-206

3 BURNHAM, W H Individual differences in the imagination of children *Ped Sem*, 1892, **2**, 204-225

4 COOLEY, C H Human nature and the social order New York Scribner's, 1922 (See esp p 88)

5 GREEN, G H Psychoanalysis in the classroom New York: Putnam, 1922 Pp xix+272

6 HALL, G S, *et al* Aspects of child life and education Boston Ginn, 1907 Pp 326 (See esp Chap. "The imaginary playmate," by T L. Smith)

7 HARVEY, N A Imaginary playmates and other mental phenomena of children Ypsilanti, Mich. State Normal Coll, 1918 Pp 12+

8 KIRKPATRICK, E A Fundamentals of child study (3rd ed) New York Macmillan, 1920 Pp 385.

9 MARTIN, L J Ghosts and projection of visual images *Amer J Psychol.*, 1915, **26**, 251-257

10. MUNROE, J P. Report at the Chicago meeting of the National Educational Association in July 1893 *Ped. Sem*, 1894, **3**, 182-184

11 NORSWORTHY, N, & WHITLEY, M T Psychology of childhood New York Macmillian, 1918 Pp xix+375

12 SMITH, T L The psychology of day dreams *Amer J Psychol*, 1904, **15**, 465-488

13 TANNER, A E The child his thinking, feeling, and doing Chicago: Rand, McNally, 1915 Pp 534

14 VOSTROVSKY, C A study of imaginary companions *Education*, 1894, **15**, 383-398

15 ———— Study in reminiscence Helen, the life history of certain imaginary companions *Stud Educ*, 1896, **1**, 243

Columbia University
New York City

UNE ÉTUDE DU CAMARADE IMAGINAIRE AU MOYEN DE QUESTIONNAIRES

(Résumé)

Sept cent un étudiants universitaires, dont 393 femmes et 308 hommes, âgés de 15 à 40 ans, l'âge moyen étant de 18 à 19 ans, ont repondu à un questionnaire fait dans le but de déterminer la fréquence de l'apparence et les traits caractéristiques généraux du "camarade imaginaire" Les principaux résultats de cette étude sont les suivants

La création du camarade imaginaire semble plus frequente chez les filles que chez les garçons Le milieu et l'état social général de l'enfant qui a un camarade imaginaire ne sont pas très différents de ceux de l'enfant qui n'a pas éprouvé cet phénomène Chez les filles, l'âge auquel le camarade imaginaire se montre pour la première fois est entre 5 et 7 ans, tandis que chez les garçons, c'est généralement après l'âge de 10 ans Le sentiment de solitude et l'inspiration d'un personnage de quelque conte sont les causes les plus usuelles de la première apparence du camarade imaginaire Dans la plupart des cas, ce camarade imaginaire est très aimé de l'enfant, qui n'aime ni à le partager ni à le discuter Le camarade imaginaire a un nom distinct et une personnalité distincte, qui deviennent si vivants que l'enfant croit à la réalité du camarade La dernière apparence du camarade imaginaire a lieu généralement après l'âge de dix ans

HURLOCK ET BURSTEIN

EINE FRAGENBOGENUNTERSUCHUNG DES "IMAGINAREN SPIELGEFAHRTEN"

(Referat)

Ein Fragebogen zur Bestimmung der Haufigkeit und der allgemeinen Eigenschaften des "imaginaren Spielgefahrten" wurde von 701 Studenten und Studentinnen (308 Mannern und 393 Frauen) beantwortet, die zwischen 15 und 40 Jahre alt waren und deren durchschnittliches Alter 18-19 Jahre war. Die Hauptbefunde aus dieser Untersuchung waren folgende:

Die Erschaffung eines imaginaren Spielgenossen scheint ofter unter Madchen als unter Knaben zu geschehen Die Umgebung und der allgemeine Hintergrund des Kindes, das einen imaginaren Speilgefahrten besitzt unterscheiden sich nicht wesentlich von denen des Kindes, das eine solche Erscheinung nicht erfahren hat Bei Madchen liegt das Alter, worin der imaginare Spielgefahrte zum ersten Mal erscheint, zwischen 5 und 7 Jahren, bei Knaben gewohnlich nach 10 Jahren Die Einsamkeit und die Begeisterung (inspiration) durch einen Karakter aus einem Geschichtenbuch stellen die haufigsten Grunde der ersten Erscheinung des imaginaren Spielgefahrten dar In den meisten Fallen wird dieser imaginare Kamerad von dem Kinde sehr hoch geschatzt, und mit anderen nicht gerne geteilt oder besprochen. Der imaginare Spielgefahrte hat einen besonderen Namen und eine bestimmte Personlichkeit, die so lebendig wird, dass das Kind an der Wirklichkeit des Spielgefahrten glaubt Die letzte Erscheinung des imaginaren Spielgefährten findet gewohnlich statt wenn das Kind mehr als zehn Jahre alt ist.

HURLOCK UND BURSTEIN

SOME EXPERIMENTS WITH KITTENS ON THE SIMPLE ALTERNATION PROBLEM*[1]

From the Animal Laboratory of the Department of Psychology of Columbia University

AUDREY M SHUEY

Up to the present time no study on simple alternation has been performed on cats In fact, only a few studies have been made on animals of any sort. Carr tested 8 white rats on simple alternation using a T-shaped discrimination box The number of trials given daily varied from 2 to 18 On the first trial every day the path from the left exit was blocked, the blockings thereafter being changed in succeeding trials. When the interval between trials was 16.5 seconds in length the animals were concerned with the act of eating and usually maintained a relatively constant position They were able accordingly to make a record of 85% correct in from 168 to 588 trials When the interval between trials was lengthened to 50 or more seconds and the rats were free to run around on an adjacent table after having been allowed a few bites of food, the animals were able to approximate a record of only 70% correct

Hunter gave 7 rats 10 trials daily on a discrimination box similar to the one used by Carr As in the latter's experiment the alternative pathways were blocked from trial to trial Twelve-second intervals were allowed between trials for eating, during which periods a fixed bodily orientation was usually maintained. Final success in the problem was reached when the animals could make a score of 87.5% correct in 4 successive days with no day's record below 80% All of Hunter's rats attained the norm in from 10 to 180 trials

The present experiments were performed on the Jenkins' problem box which was used in a previous experiment reported by the writer (4). A detailed description of the apparatus has been given in the former study and in an earlier article by Jenkins (3), hence only the

*Accepted for publication by C J Warden of the Editorial Board and received in the Editorial Office, March 28, 1932

[1]The experiments were performed in the Animal Laboratory of the Department of Psychology of Columbia University under the direction of Professor C J Warden

main features need be given here The test cage, the ground plan of which is shown in Figure 1, was 135 centimeters in diameter. The diameter of the inner cage was 45 centimeters The walls of both cages were 45 centimeters high and were made of ½-inch wire mesh. On the floor of the test cage were three plates, 1, 2, 3, each 15 centimeters in diameter and spaced as shown in Figure 1. These plates were raised 1 centimeter above the floor and were of the same material and color as the floor. They were strung with fine copper wires so that electric shock could be used when desired. A one-way light screen surrounding the box excluded the experimenter from the view of the subject, the cage itself being lighted by a 100-watt Mazda bulb above the center of the box

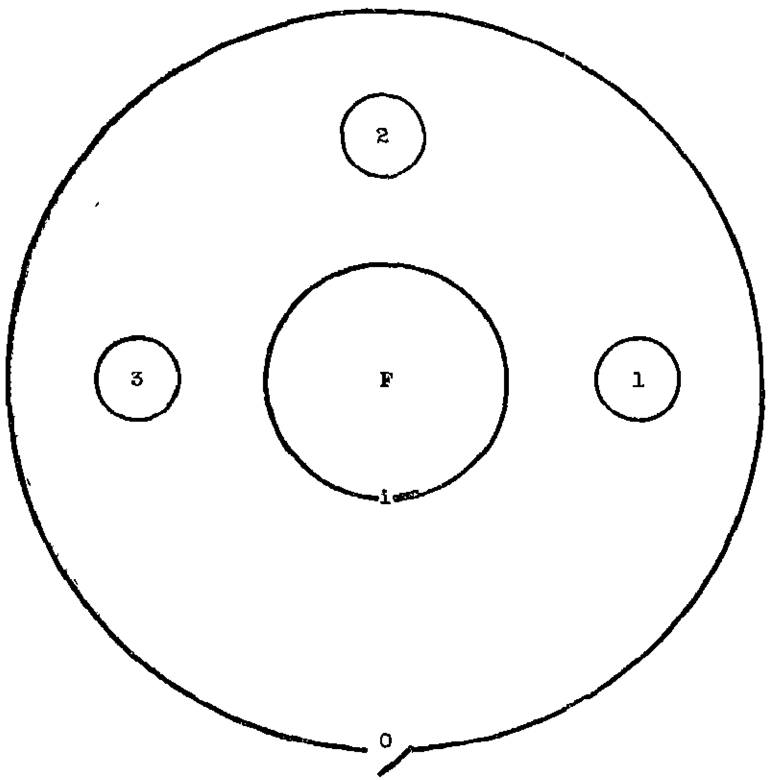


FIGURE 1

The subjects of the experiment included 4 males and 5 females of the 82 kittens observed in the writer's study already mentioned. These kittens, the males numbered 20, 21, 22, 23, and the females numbered 66, 67, 68, 69, and 70, were all between 11 and 12 weeks of age at the beginning of the training on the alternation problem When they were started on this problem, the 9 kittens had all reached the point referred to in the foregoing study as the advanced problems of the previous study (4). This means that all of the subjects used in the present experiment had learned to touch any one plate, then had learned to touch any two plates, and then all three of the plates Each kitten had learned to make this last response up to a point of 9 perfect trials out of 10 Therefore, when they were introduced to the alternation problem, the animals were accustomed to reacting to all three plates

The general method of procedure was simple. The outer door, *o*, the means of entrance to the compartment, was opened and closed manually by the experimenter The inner door, *i*, was held shut or was allowed to slide back automatically by the operation of an electromagnet When this door was opened the animal in the outer part was given free access to the incentive in the inner cage Before the inner door was released the kitten had to react by touching one of the release plates as determined by the experimenter.

Bits of raw beef and fresh or canned milk were the incentives All kittens were given 12 trials daily, 6 in the morning and 6 in the evening, including Sundays, between the hours of 10 and 1 All animals were fed in the living quarters immediately following experimentation, whereupon food was removed from them until after the subsequent test period All subjects were in good health throughout the running of the tests

Since somewhat different techniques were used in the earlier and later work, it will be convenient to discuss the results of each separately

Experiment I

In the first experiment although all three plates were available it was only necessary for the kitten to touch Plates 1 or 3. If the kitten happened to touch Plate 1 on the very first trial it was to touch Plate 1 on Trials 3, 5, 7, and all the odd-numbered trials thereafter, and touch Plate 3 on the even-numbered trials If a kitten touched Plate 3 on the first trial it was thereafter to begin the daily series by reacting

to Plate 3, alternating on succeeding trials. Whenever a kitten touched the wrong plate it was given a mild shock, one that proved to be sufficient as determined by earlier tests upon other kittens. For shocking purposes, the e.m.f. was set at 400 volts with a current-reading of 0.04 m a Unless the kitten actually touched the wrong plate the trial was considered correct. It required, on the average, about 5 seconds for the kitten to touch the plate and secure the reward, and there was approximately 10 seconds between successive trials. When the experimenter put the kittens into the problem box she placed them in such a position that their heads consistently pointed toward the knob of the inner door.

It will be noted that the alternation set-up as used in this experiment differs from the regular alternation situation as used by Carr and Hunter In their studies it was wholly a matter of choice of one of two directions In the present experiment the animal must not only choose the path, but also must afterward step on a proper plate and then turn back to the inner door. However, as the kittens had already formed a connection between the stepping on a plate and the opening of the door to the food box, it did not seem that the task set was fundamentally different from the ordinary alternation set-up

When the kittens were tested, however, we noted that they tended to go all the way around the inner cage in each trial. If the door did not open when it touched one plate, the kitten would sometimes turn back; but more often it would go on around the box touching one of the other plates on the way This behavior was particularly true of kittens No 21 and No 66. The latter kitten would usually take a path similar to the following past 1, 2, and touching 3, or past 1, 2, touching 3 and 1, or past 3, 2, touching 1; or past 3, 2, touching 1 and 3. Another would go back and forth between Plates 1 and 3, by Plate 2, i.e., past 1, 2, touching 3, 2, 1, or past 1, 2, touching 3, past 2, touching 1.

Of the four animals tested, not one succeeded in mastering the problem in the 432 to 1200 trials given them. As will be noted from Table 1, two of the kittens that were given 1200 trials were successful in only 49 1 and 49.9%, respectively, of the total number of trials given, while the third made an average of 53 0% correct Kittens No 20 and No 66 made no improvement in successive quarters of the total number of trials, while kitten No 21 made almost negligible progress Table 2 shows the best records attained by the animals throughout the experiment. It will be observed that two of

TABLE 1
SHOWING THE PERCENTAGE OF PERFECT TRIALS IN EACH SUCCESSIVE 48, WITH SUMMARY, FOR EACH KITTEN IN EXPERIMENT I

	Animals			
Trials	No 20	No 21	No 22	No. 66
1-48	50 0%	47 9%	43.8%	45 8%
49-96	47 9	39 6	54 2	52 1
97-144	56 3	47 9	52 1	47 9
145-192	56 3	58 4	54 2	50 0
193-240	60 4	56 3	52 1	54 2
241-288	52 1	50 0	50 0	41 7
289-336	43 8	39 6	50 0	45 8
337-384	56 3	62 5	58 4	62 5
385-432	50 0	43.8	52 1*	45 8
433-480	47 9	41 7		50 0
481-528	60 4	27 1		41 7
529-576	60 4	47 9		39 6
577-624	58 4	66 7		50 0
625-672	43 8	50 0		52 1
673-720	52 1	41.7		41 7
721-768	52 1	45 8		52 1
769-816	60 4	50 0		58 4
817-864	56 3	45 8		45 8
865-912	66 7	60 4		58 4
913-960	50 0	58 4		54 2
961-1008	50 0	54 2		47 9
1009-1056	47 9	52 1		54 2
1057-1104	52.1	58 4		39 6
1105-1152	39 6	54 2		47 9
1153-1200	52 1	52 1		50 0
1-288	53 8%	50 0%	51 1%	48 6%
289-576	53 1	43 7		47 6
577-864	53 8	50 0		50 0
865-1200	51 2	55 7		50 3
Av	53 0%	49 9%		49 1%

*Animal No 22 died at the end of Trial 432

the kittens made records of 11 out of 12 successive perfect trials or a score of 91 7%, that these same two scored 79.7% in 24 trials, that one of them scored 72 2% in 36 successive trials, and that the other attained a record of 68 8% over a period of 48 trials This latter score was considerably below the standard required by Hunter and Carr.

The results of Table 3 show that kitten No 20 had a preference for Plate 1 throughout, that kitten No 66 had a preference for Plate 3, and that kitten No 21 started with a preference for Plate 1 but

gradually went more and more often to Plate 3. Of these kittens, Nos. 20 and 66 were required to react to Plate 3 first in each series of trials, while kittens Nos. 21 and 22 were required to touch Plate 1 first.

TABLE 2

BEST RECORDS ATTAINED IN EXPERIMENT I*

Norm series	Animals No. 20	No. 21	No. 22	No. 66
12 tr. in suc.	91.7% (861)	91.7% (603)	66.7% (326,382, 401)	75.0% (341,659, 807,905)
24 tr. in suc.	79.2 (186)	79.2 (587)	58.3 (326)	70.8 (329)
36 tr. in suc.	69.4 (186,861)	72.2 (580)	58.3 (326)	61.1 (329,863, 1013)
48 tr. in suc.	68.8 (469)	64.6 (580)	54.2 (326)	62.5 (329)
60 tr. in suc.	65.0 (469)	63.3 (580)	50.0 (326)	61.7 (329)

*The percentages indicate the best record for each norm throughout the experiment. The stage of learning when each record was made varies for each kitten and is shown in terms of number of trials after each percentage.

TABLE 3

SHOWING PREFERENCES FOR PLATES 1 OR 3 IN EXPERIMENT I

Trials	Animals No. 20		No. 21		No. 22		No. 66	
	1	3	1	3	1	3	1	3
1-288	64.6%	43.1%	72.2%	27.8%	73.8%	28.4%	19.5%	77.8%
289-576	55.5	50.7	69.4	18.1			54.2	41.0
577-864	61.1	46.5	27.1	72.8			21.5	78.5
865-1200	56.5	45.8	32.1	79.2			26.8	73.8
Av.	59.4%	46.5%	50.2%	49.5%			30.5%	67.8%

TABLE 4

SHOWING THE PERCENTAGE OF PERFECT INITIAL CHOICES FOR EACH KITTEN IN EXPERIMENT I

Trials	Animals No. 20	No. 21	No. 22	No. 66
1-288	45.8%	79.2%	75.0%	66.7%
289-576	56.3	66.7		37.5
577-864	47.9	25.8		83.3
865-1200	53.6	46.4		78.6
Av.	50.9%	54.5%		66.5%

It will be seen from the data of Table 4 that mastery of the initial trial of each series is no indication that the alternation habit has become established, for the two kittens that made the highest scores in the test showed little or no mastery of the first trial of the series. On the other hand, the poorest of the kittens, No 66, established the correct habit of going first to Plate 3, but continued to go to Plate 3 thereafter, instead of alternating between Plates 3 and 1.

Experiment II

It seemed possible that the failure to learn the alternation habit in the above experiment might be due in part to the specific set-up The presence of Plate 2 in particular might have been confusing to the animal In order to avoid this difficulty a wire screen, wedge-shaped, extending from the inner to the outer cages and from the top to the bottom of the box completely cut off Plate 2 and made the problem box into a 2-way circular maze. Since this double partition prevented the animals from running around the box, we seemed to have a more nearly typical alternation arrangement

The 5 kittens trained in Experiment II were given from 240 to 1500 trials, all experimental procedure and technique being the same as for the animals in Experiment I These 5 animals had been trained to react to Plates 1, 2, and 3, as had the animals in the first experiment reported.

In general, somewhat better records were obtained in Experiment II than in Experiment I. However, even in the second experiment the majority of the kittens failed to master the problem. Of the 4 kittens that were given more than 1100 trials, two showed no improvement in successive quarters of trials. Of the total trials, only 50 4% and 51.4%, respectively, were performed correctly The best records made by these two kittens, No 69 and No 70, were even lower than some of those obtained in the first experiment As may be observed from Table 6, the highest record these subjects made for 24 successive trials was 70.8% For longer series of trials the highest percentages of perfect scores ranged from 58 3 to 69.4. These scores are probably too low to signify any learning whatsoever.

The third kitten, No 68, made a somewhat larger percentage of perfect trials out of the total number, as may be observed from Table 5 This subject likewise made, on the whole, superior records to those of any kitten in Experiment I, having on different occasions made 11 out of 12 trials perfect, also 19 out of 24, 24 out of 36, 34

out of 48, and 41 out of 60, or, respectively, in terms of percentages, 91.7, 79.2, 75, 70.8, and 68.3. While records for this animal are not altogether reassuring, there seems to be some slight evidence of an imperfectly formed alternation habit.

TABLE 5
SHOWING THE PERCENTAGE OF PERFECT TRIALS IN EACH SUCCESSIVE 48, WITH SUMMARY, FOR EACH KITTEN IN EXPERIMENT II

	Animals				
Trials	No. 23	No. 67	No. 68	No. 69	No. 70
1-48	39.6%	54.2%	62.5%	45.8%	50.0%
49-96	54.2	62.5	58.4	50.0	45.8
97-144	50.0	52.1	50.0	54.2	52.1
145-192	47.9	58.4	52.1	60.4	60.4
193-240	47.9	64.6*	45.8	47.9	41.7
241-288	43.8		37.5	41.7	56.3
289-336	52.1		56.3	50.0	50.0
337-384	50.0		52.1	50.0	50.0
385-432	50.0		54.2	56.3	41.7
433-480	58.4		52.1	56.3	54.2
481-528	62.5		68.8	56.3	58.3
529-576	62.5		56.3	47.9	47.9
577-624	60.4		56.3	56.3	64.6
625-672	79.2		54.2	54.2	43.8
673-720	58.4		45.8	39.6	52.1
721-768	50.0		54.2	54.2	41.7
769-816	52.1		54.2	60.4	37.5
817-864	50.0		68.8	52.1	56.3
865-912	58.4		62.5	47.9	64.6
913-960	64.6		47.9	39.6	52.1
961-1008	60.4		58.4	58.4	45.8
1009-1056	66.7		45.8	47.9	41.7
1057-1104	68.8		64.6	54.2	56.3
1105-1152	79.2		47.9		52.1
1153-1200	52.1†		50.0		47.9
1201-1248	52.1		52.1		52.1
1249-1296	62.5		58.4		52.1
1297-1344	66.7		60.4		37.5
1345-1392	62.5		56.3		43.8
1393-1440	54.2		50.0		56.3
1441-1488	56.3		54.2		56.3
1489-1500	66.7		58.4		52.1
1-384	48.2%		51.8%	50.0%	50.8%
385-768	60.2		55.2	52.6	50.5
769-1152	62.5		56.3	51.5	50.8
1153-1500	58.4		54.6		49.5
Av.	57.3%		54.5%	51.4%	50.4%

*Animal No. 67 died at the end of Trial 240.
†Twelve days' interval between Trials 1170 and 1171 of kitten No. 23.

TABLE 6

BEST RECORDS ATTAINED IN EXPERIMENT II*

Norm series	Animals No 23	No 68	No 69	No 70
12 tr in suc.	91 7% (643,661,900,1099,1141)	91 7% (1146)	83 3% (175,343,521,655,803,962)	91 7% (593)
24 tr in suc	83 3 (631,643,895)	79 2 (1053)	70 8 (163)	70 8 (469)
36 tr in suc	83 3 (643)	75 0 (493)	63 9 (151)	69 4 (571)
48 tr in suc	81 3 (631)	70 8 (493)	58 3 (139,772)	66 7 (571)
60 tr in suc	80 0 (631)	68 3 (493)	60 0 (127)	63 3 (571,865)

*The percentages indicate the best record for each norm throughout the experiment The stage of learning when each record was made varies for each kitten and is shown in terms of number of trials after each percentage

TABLE 7

SHOWING PREFERENCES FOR PLATES 1 OR 3 IN EXPERIMENT II

Trials	Animals No 23		No 68		No 69		No 70	
	1	3	1	3	1	3	1	3
1-384	69 3%	27 1%	56 3%	47 4%	38 0%	61.9%	58 3%	43.2%
385-768	81 8	38 5	66 7	43 8	65 6	39 6	49 5	51 6
769-1152	75 0	50 0	64 6	47 9	62 5	40 5	72 4	29 2
1153-1500	63 2	53 5	74 1	35 1			80 5	18 4
Av	72 3%	42 3%	65 4%	43 5%	55 4%	47 3%	65 2%	35 6%

Only one kitten appears to have definitely established the alternation habit This kitten, No 23, made 57 3% perfect responses out of a total of 1500, the score for the last 1116 trials being slightly over 60%. It is possible that this animal might have made a higher score if there had not been a 12-day interval separating Trials 1152 and 1153. In spite of this fact, the kitten reached relatively high norms on several occasions, as may be seen from Table 6 He attained a norm of 11 out of 12 trials (91 7%) five times, and made a score of 83 3% in 36 trials covering 3 days' testing, with no day's score below 75%. Furthermore, this same kitten scored 80% correct in 60 successive trials, i e, 5 days' testing, with no day's score below 75% These latter records are almost as high as those secured by Hunter and Carr Furthermore, if momentarily one disregards the main requirement, that of stepping on Plate 1 or Plate 3, and notes only the direction the animal first took at the beginning of every trial, he finds that this kitten scored 86.1% out of 36 successive trials, 83 1% out of 48 trials, and 81.7% out of 60 trials When the data for the other three animals were treated similarly, their best records were not changed as in the present case.

All of the kittens in Experiment II showed a preference for going to the right rather than to the left, since they went more often to Plate 1 than to Plate 3 This was true in spite of the fact that all but one, No. 70, were required to begin each series of trials with Plate 3 The general tendency to go to the right rather than to the left was observed in these same kittens in trials preliminary to these on the alternation problem

Since all kittens favored the right, or Plate 1, it would be expected that only kitten No. 70, the one that was required to touch Plate I first in each series of 6 trials, would show a large percentage of perfect

TABLE 8
SHOWING THE PERCENTAGE OF PERFECT INITIAL CHOICES FOR EACH KITTEN IN EXPERIMENT II

	Animals			
Trials	No 23	No 68	No 69	No 70
1-384	18 8%	53 1%	68 8%	54.7%
385-768	40 6	68 8	46.9	54 7
769-1152	62 5	53 1	42 9	65 6
1153-1500	70 3	39 1		70 3
Av	48 1%	53 5%	52 9%	61 3%

initial choices Table 8 shows this to be true A comparison of Tables 7 and 8 indicate that, while the percentages of perfect initial choices by kittens No 23, No. 68, and No. 69, were low, yet they were higher than the total percentages of times Plate 3 was touched correctly That is, the less favored plate was more likely to be touched in the first series of 6 trials than in the succeeding trials It is interesting, finally, to note that the best kitten in the last quarter of the 1500 trials made a perfect score in 70 3% of the initial choices, while in the first quarter the corresponding value was only 18 8%.

It should be pointed out that after every trial the kitten was lifted out of the food compartment through the top of the inner cage and placed on the top of the outer cage while the experimenter placed a bit of meat in the center and closed the door for the next trial Although these operations required only about 10 seconds, the animals were observed to wander about on the top of the box and thus loose their orientation from trial to trial As we have noted before, Carr found that when he removed his rats to a nearby table top in between trials for 50 or more seconds the percentage of correct choices was decreased from 85 to 70% Probably our animals would have made higher scores if they had been transferred directly from the food box to the entrance box. It has been noted that better records were obtained under the conditions of the experiment where the animals were prevented from running around the box during the trials, But even here the animals were required not merely to go to the right or to the left, but to touch a plate in the appropriate segment and then to turn back to the food box This additional requirement seems to have added something to the difficulty of the problem as compared with the usual simple-alternation technique

Summary

1 After 9 kittens had learned to step on all three plates of the Jenkins' problem box, an attempt was made to establish in them the habit of reacting alternately to the right and left plates

2 In Experiment I, which permitted access to the unused Plate 2, little evidence of a consistent habit appeared, although from 432 to 1200 trials were given to the animals. The best record over a period of 36 trials was 72 2%, attained by the kitten on Trial 580

3 In Experiment II, in which access to Plate 2 was prevented by a screen, the results were somewhat better where 1100 to 1500 trials were given A fairly consistent alternation habit was established in

one kitten, the best record being 83.3% for Trials 643 through 678 and 80% for Trials 631 through 690.

References

1. Carr, H. A. The alternation problem. *J. Anim. Behav.*, 1917, 7, 365-384.
2. Hunter, W. S. Kinaesthetic sensory processes in the white rat. *Psychol. Bull.*, 1918, 15, 36-37.
3. Jenkins, T. N. A standard problem box of multiple complexity for use in comparative studies. *J. Comp. Psychol.*, 1927, 7, 129-144.
4. Shuey, A. M. The limits of learning ability in kittens. *Genet. Psychol. Monog.*, 1931, 10, 287-378.

Columbia University
New York City

QUELQUES EXPÉRIENCES AVEC DES PETITS CHATS SUR LE PROBLÈME D'ALTERNATION SIMPLE

(Résume)

Après que neuf petits chats, âgés de 11 à 12 semaines, avaient appris à mettre le pied sur toutes les trois plaques de la boîte à problèmes de Jenkins, on a essayé de faire acquérir aux chats l'habitude de réagir dans les épreuves alternées aux plaques droite et gauche. On a fait douze épreuves par jour, 6 le matin et 6 le soir, avec du boeuf et du lait comme récompense et un choc comme punition. On s'est servi de quatre des animaux dans la première expérience où l'on a permis l'accès à la plaque du centre. Aucun animal de ce groupe n'a donné un meilleur resultat que celui de 72,2 pour cent dans une periode de 36 épreuves bien que l'on ait fait subir de 432 à 1200 épreuves à chacun. On a fait subir de 1100 a 1500 épreuves à cinq animaux dans une seconde expérience où l'accès à la Plaque 2 a été empêché par un ecran de fil de fer en forme de coin s'étendant des cages intérieures aux cages extérieures. On a etabli une habitude assez constante d'alternation chez un petit chat qui a donné un resultat de 83,3 pour cent correctes sur une période de 36 épreuves et un résultat de 80 pour cent correctes sur une période de 60 épreuves.

Shuey

EINIGE VERSUCHE AN KATZCHEN MIT DER EINFACHEN ABWECHSLUNGSAUFGABE

(Referat)

Nachdem neun Katzchen, 11 bis 12 Wochen alt, gelernt hatten, auf alle drei Platten der Aufgabenkiste (problem box) von Jenkins zu treten, versuchten wir, in ihnen die Gewohnheit zu formen, in abwechselnden Proben (trials) abwechselnd auf die rechte und die linke Platte zu treten. Es wurden taglich zwolf Proben gemacht, sechs morgens und sechs abends; als Lohn dienten Rindfleisch und Milch, und als Strafe diente elektrischer Reiz. In dem ersten Versuch, in dem die mittlere Platte betreten werden durfte, wurden vier von den Tiere verwendet. Kein Tier aus dieser Gruppe erzielte eine bessere Leistung (record) als 72.2% wahrend einer Periode, die sich uber 36 Proben erstreckte, obwohl an jedem Tier 432 bis 1200 Proben gemacht wurden. In einem zweiten Versuch wurde an funf Tieren 1100 bis 1200 Proben gemacht. In diesem Versuch wurde der Zugang zu der zweiten Platte mit einer keilformigen Drahtmaschenschutzwand (wire mesh screen), die sich von den inneren zu den ausseren Kafigen erstreckte, versperrt. In einem Katzchen bildete sich eine ziemlich bestandige Gewohnheit der Abwechslung (alternation habit). Dieses Katzchen erzielte in 36 Versuchen eine Leistung von 83.3% an richtigen Wahlen, und wahrend einer Periode, die sich uber 60 Versuchen erstreckte, eine Leistung von 80%.

SHUEY

LEARNING VERSUS SKILL IN RATS*†

From the Anthropoid Experiment Station of Yale University, Orange Park, Florida

JOSEPH G. YOSHIOKA

Learning experiments with animals in general are beset at the start with one outstanding difficulty, namely, that we must let them know what to do, but have no adequate means to do so Verbal instruction being out of the question, we resort to a rather crude method of putting them in a given situation and letting them find out for themselves the nature of the problem and its solution, trusting to luck that they will do so Under such a circumstance, there is no assurance, if a solution is made, as to whether it has been reached deliberately or fortuitously. If a solution is not achieved, it is equally doubtful as to whether the failure lies in the difficulty of understanding what is to be done or in solving the problem If, however, the solution is repeated consistently in subsequent trials, it is assumed that the animals have learned the problem This empirical practice seems to have surmounted the obstacle that learning experiments with animals must be carried out without instruction, but in reality it has not a whit obviated the intrinsic difficulty pointed out above, as is shown by divergent opinions and practices in selecting "learning scores" and by heated discussion about validity and reliability of maze learning

Some time ago the writer proposed a view (2) that maze learning consists at least of two phases, learning proper and acquisition of skill, and implied that the usual maze scores represented more skill than learning. Maze scores taken with the usual criteria, however, cannot easily be analyzed into learning and skill scores because of our lack of knowledge of the essential nature of maze learning. As Hull (1) says, maze learning is not a relatively simple process as assumed by some, but in reality one of great complexity In order to elucidate further the learning and skill phases of performance by rats, a simple learning experiment was performed, wherein the elements in-

*Accepted by Calvin P Stone of the Editorial Board and received in the Editorial Office, February 29, 1932

†Studies from the Behavior Research Fund, Chicago Series B, No 183

volved in learning could be specifically stated and the mastery of these elements could be scored without ambiguity The following paragraphs will describe the procedure and results

Method of the Present Experiment

Thirty-three albino and pied rats, about 4 months old, mostly males, selected at random from our stock colony, were put to a simple task of eating 10 large-sized sunflower seeds in 10 minutes in a new situation The animals had been reared in the stock-room, had never before been taken out of the room, nor been used in any experiment, and had eaten no other food than milk and the Steenbock diet Starting at 10 A M daily they were taken to the experimental room in groups of five or six and placed individually within an enclosure standing on a linoleum floor and surrounded by a sheet-iron wall, 14 inches square and 12 inches high. Ten sunflower seeds, about 14 mm. long, 8 mm wide, 5 mm thick, were piled up at the center of the floor within the enclosure Two such enclosures were used so that two animals could be tested simultaneously After 10 minutes the animals were taken out, and a record was made as to the number of seeds eaten, the number of husk-fragments, the number of feces, if any, and whether traces of urine were present In case an animal ate 10 seeds in less than 10 minutes he was taken out as soon as the last seed was finished, and the time was noted. This daily test continued for 15 days.

Experimental Data

Learning in this experiment involved habituation to the new situation and consumption of a definite quantity of new food within a set time limit. It was observed that on the first day all the animals were very much upset in the enclosure, 27 out of the total 33 defecated, 2 urinated, some explored frantically, trying to escape, and others sat immobile because of fear. None paid any attention to the seeds at first, and although 11 of them later nibbled at them, no rat succeeded in eating 10 seeds. On the second day, 10 animals attacked the seeds, 2 of them finished 10 seeds, and the number of rats defecating dropped to 13 The degree of habituation to the new situation (enclosure) was roughly proportional inversely to their emotional excitement expressed behaviorially by defecation, and so was scored in terms of defecation. Table 1 shows the distribution of rats defecating and urinating and the frequency of defecations on succeeding tests.

TABLE 1

Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
No of rats defecating	27	13	11	11	9	7	6	5	5	5	6	2	2	2	2
No of feces	82	39	39	23	30	19	14	12	12	9	18	6	4	4	4
No of rats urinating	2	4	1	1	1	2	3	1	3	2	0	2	0	1	0

The number of the rats defecating and the number of feces per trial are distributed in the manner of the usual error or time scores, i e , starting high and dropping down markedly in the first few days, and thereafter gradually levelling off. The number of rats urinating, however, showed no such tendency. Two or three exceptional rats urinated daily throughout the period The number of feces per rat, that is, the quotient obtained by dividing the second row by the first row of numbers in Table 1, failed likewise to show a progressive degree of habituation Exploratory movement in the enclosure was found too inconstant to be a measure of habituation, for some rats moved about freely from fear, others from familiarity, likewise, some rats sat still from fear, others from complacency In general, habituation in a new situation by the animals is inversely proportional to emotional excitement, as mentioned above, and may be scored objectively *by the presence or absence of defecation.* To prove further that defecation indicates emotional excitement and hence less habituation, a relationship between defecation and the first attack on the seeds (more habituation) was sought and expressed in terms of Yule's (3, pp. 25-39) coefficient of association. From the data obtained on the first day each rat was scored for defecation and for eating the seeds The coefficient of association between these two scores for the group was found to be —.41 Many rats that defecated did not eat the seeds, while others that ate the seeds did not defecate In this connection, it may be of interest to mention another coefficient of association in behavior between two different situations. Twenty of these rats were used later for another experiment and ran a diamond maze From the data obtained on the first day in the enclosure and the corresponding data in the maze each rat was scored for sitting in these two situations The coefficient was —26, and insignificant. The rats that remained immobile in the enclosure did

not necessarily behave likewise in the maze. Others that were active in the enclosure were not necessarily so active in the maze. It seems that each new situation may excite the same rats differently and that no prediction can be made on the excitability of an animal in a new situation from its behavior in a previous situation Emotional excitability is generally ignored in learning scores, but, since there is a slight negative association between excitability and learning, as shown above, emotional scores such as can be obtained from the frequency of defecation may be of use in weighting individual scores in a learning experiment in which finer individual differences are sought For example, a rat that scored the same in maze learning as another, but had had less practice in perfecting the learning than the latter, because of its higher excitability, and, consequently, of longer time needed for habituation, can have its learning score made higher by weighting in reference to its emotional score.

The next measure of learning considered pertains to eating the sunflower seeds. Table 2 gives the learning scores obtained with different criteria The first row of Table 2 gives the distribution of rats that began to eat the seeds for the first time on different days On the first day one-third of the group attacked the seeds, and thereafter one or two rats were added daily to the list of learners, and at the end of the experiment only four rats remained unresponsive This distribution ranks the individuals in terms of mastery of this phase of learning, but is relatively unstable because some rats that ate the seeds on the first day did not necessarily keep on eating on the subsequent days A more stable distribution for group comparison is given in the second row of Table 2, where the number of rats eating seeds on a particular day is tabulated. The number of learners increased gradually from 11 on the first day to 26 on the last day. Figure 1 shows the curve denoting the distribution of learners.

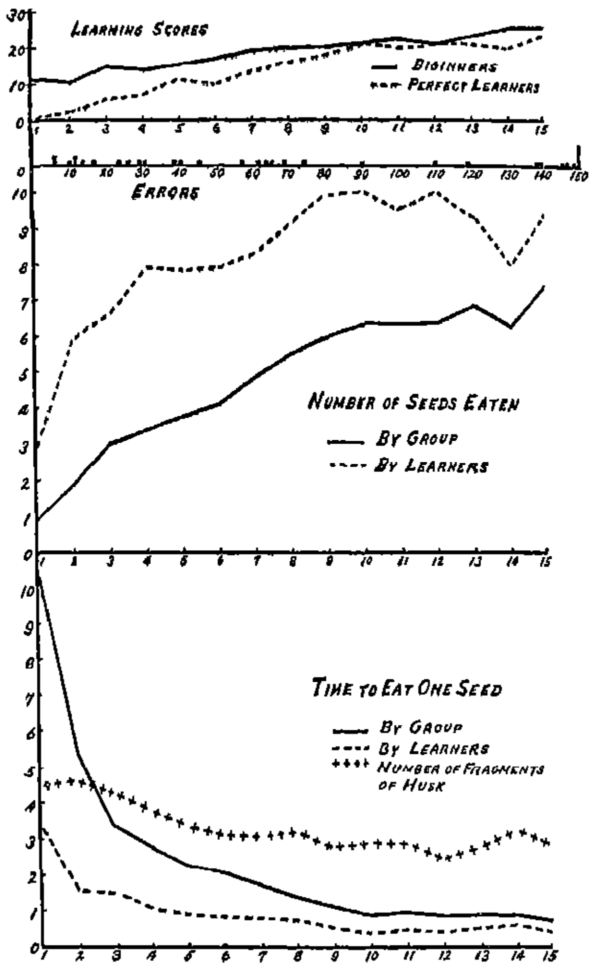
Next we may consider learning as measured by meeting the time-limit requirement, that is, to eat 10 seeds in 10 minutes. The third row of Table 2 gives the distribution of rats fulfilling the requirement. On the first day no rat met the criterion, from the second day to the tenth day the number of perfect learners increased steadily, and thereafter there was no significant increase, the curve as plotted in Figure 1 (broken line) showing a plateau at this stage, and at the end of the experiment seven had failed to meet the criterion This distribution represents the group performance scored in refer-

FIGURE 1

FIGURE 2

FIGURE 3

FIGURE 4



ence to all the criteria set for the learning, and is to be used for group comparison The individual ranking was obtained by subtracting the number of perfect learners on a particular day from that on the succeeding day, and is shown in the fourth row of Table 2. This distribution represents the individual differences in meeting the learning criteria The negative number on the 6th, 11th, and

TABLE 2
LEARNING SCORES
$N = 33$

Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Rats beginning to eat seeds	11	2	3	0	2	1	2	1	0	1	1	0	2	3	0
All the rats eating seeds	11	10	15	14	16	17	19	20	20	21	22	21	24	26	26
Perfect learners	0	2	6	7	11	10	13	16	18	21	20	21	21	20	24
Rats reaching perfect learning	0	2	4	1	4	—1	3	3	2	3	—1	1	0	—1	4
Average number of seeds eaten	.97	1.79	3.00	3.36	3.76	4.06	4.82	5.52	6.00	6.36	6.33	6.36	6.83	6.30	7.39
Average number of seeds eaten by learners	2.91	5.90	6.60	7.93	7.75	7.88	8.26	9.10	9.90	10.00	9.50	10.00	9.38	8.00	9.38
Average time to eat one seed*	10.30	5.52	3.33	2.70	2.33	2.05	1.71	1.40	1.13	94	1.00	93	94	92	71
Average time to eat one seed by learners*	3.44	1.63	1.52	1.07	97	84	84	69	52	39	47	36	50	63	42
Average number of husk-fragments	4.53	4.73	4.29	3.83	3.41	3.10	3.04	3.22	2.75	2.87	2.77	2.54	2.76	3.33	2.89

*Time is given in minutes

14th days means that one rat that made a perfect score on the previous day failed to do so. By making the criteria more difficult to meet, say, eating 10 seeds in 10 minutes for 2 successive days, the negative number can be eliminated. For still finer individual differentiation the steps that each individual had taken before reaching the stage of perfect learning were considered, and an error score was assigned to each as follows: for each seed left uneaten one point was given as an error. If none of the seeds was eaten throughout the learning period, the error score would amount to 150. From this total one point was subtracted for each seed eaten so that a rat that had eaten the greatest number of seeds would have the minimum error score. After reaching the stage of perfect learning, the individual error score does not change since the daily error score will be zero. Hence, errors thus scored rank the individuals in terms of relative difficulty for perfecting the learning. Figure 2 gives the individual distribution in terms of errors. The piling-up of individuals on the last score was due to the fact that five rats had not yet learned perfectly at the end of the experiment. An extension of the test for a few more days would have differentiated these backward animals.

A more conventional way of plotting a learning curve is in terms of some phases of performance, and to one of these we now direct our attention. Learning in this experiment can be expressed in terms of the number per rat of seeds eaten daily. This number is shown in the fifth row of Table 2, and is plotted in Figure 3. The curve shows the typical characteristics of a learning curve. It is noted, however, that the averages of seeds eaten daily at the plateau were between 6 and 7, but, as shown above, at this stage about one-third of the group were eating 10 seeds daily, having reached the maximum learning. The low group averages were caused by zero scores of some rats that had not yet begun to eat a single seed. If learning is to be expressed by the average number per rat of seeds eaten daily, zero scores should be excluded since we want to show here the process of learning from the point when the rats began to eat seeds to the final stage when they learned to eat 10 seeds. This was done; after excluding the zero scores the distribution became as shown in the sixth row of Table 2 and as plotted in Figure 3 (broken line). Here in the plateau stage the average number per rat of seeds eaten daily was between 9 and 10. This is more in agreement with the result obtained from the distribution of perfect learners.

Learning in this experiment can also be expressed in terms of average time consumed per rat to eat one seed Such is shown in the seventh row of Table 2, and is plotted in Figure 4. The curve conforms more or less to a typical learning curve Attention must be called, however, to the fact that from the 10th to the 15th day, excepting the 11th day, the average time per rat to eat one seed was less than 1 minute. The learning criterion required the rat to eat 10 seeds in 10 minutes, or approximately one seed in 1 minute Here the performance overshot the criterion, and hence should be regarded as representing skill, and not learning The objection to including zero scores in computing the average number per rat of seeds eaten can be raised here as much as to the similar inclusion in computing the average time to eat one seed. A rat that did not eat a single seed in 10 minutes could not be given any time score because the time was indeterminate, not zero. The average time given above was obtained by dividing the total number of seeds eaten on a particular day by the total time spent and by dividing the quotient again by the number of rats in the group Thus the non-eating rats were given zero for time score, which was both impossible and absurd When the zero scores were excluded the distribution took the form given in the eighth row of Table 2, and plotted in Figure 4 (broken line) Here, after the sixth day the average time per rat to eat one seed was less than 1 minute, indicating an early onset of skill.

Skill in this experiment can be shown in another way The number of husk-fragments of seeds eaten was averaged per rat per seed, excluding zero scores, and tabulated for consecutive days, as is shown in the ninth row of Table 2, and plotted in Figure 4 (crosses) It was observed that in the beginning the rats chewed up the husk into many fragments in opening the seed, but as they became more adapted to this food they were able to open the seed quickly by breaking the husk into a few, large fragments As shown in the curve, after the sixth day the average number of fragments was about 3 per seed Skill represented here is in perfect accord with that shown by the average time for eating one seed. It is said, therefore, that skill can be demonstrated in a learning experiment by time scores when criteria of learning are unequivocal and definitely stated, or by independent scoring for some phases of performance if such can be isolated

Summarizing, one may say that in a simple learning experiment

with rats, wherein learning involved habituation to a new situation and eating 10 sunflower seeds in 10 minutes, to which they had not been accustomed, it was shown that the degree of habituation could be scored objectively by frequency of defecation, indicative of emotional excitement, and learning proper by various scoring systems. For group comparison of learning the distribution of rats in terms of learning criteria seemed the most suitable because it represented the group performance scored in terms of the set criteria For individual differentiation of learning error scores were found to be the best because they differentiated the individuals the most, and it was suggested that these scores, errors or otherwise, could further be improved by weighting in terms of the emotional scores obtained for habituation. Learning was scored also, first, in terms of the average number per rat of seeds eaten daily, and, secondly, in terms of the average time consumed per rat to eat one seed Exclusion of zero scores from the computation of the averages in these two cases gave results more in agreement with the results obtained by the criterion method Skill was demonstrated by the time scores when the average time for eating one seed became less than 1 minute, thus the speed reaching below the time limit allowed by the criteria Skill was also demonstrated by the average number of husk-fragments per seed eaten, and these two methods showed fair agreement in the trend of skill.

In conclusion, it is said that skill can be differentiated from learning in a learning experiment with rats when the criteria of learning are unequivocal and definitely stated or when some phases of performance indicative of skill can be isolated, and that scoring for learning and the basis of such scoring should be made selectively for a particular end in view, that is, for group comparison or for gross or finer individual differentiation

References

1. HULL, C L The goal gradient hypothesis and maze learning *Psychol Rev*, 1932, **39**, 25-43
2. YOSHIOKA, J G. What is maze learning for the rat? *J Genet Psychol.*, 1929, **36**, 51-57
3. YULE, G U An introduction to the theory of statistics. Philadelphia. Lippincott, 1924 Pp 420

Anthropoid Experiment Station of Yale University
Orange Park, Florida

L'HABILETÉ ET L'APPRENTISSAGE CHEZ LE RAT

(Résumé)

Une fois par jour pendant quinze jours on a donné à trente-trois rats, âgés d'environ quatre mois, la simple tâche de manger dix semences de tournesol en dix minutes dans une situation nouvelle, où ils n'avaient jamais mangé auparavant. Le degré d'habituation à la situation nouvelle a été évalué par la fréquence de défécation pendant le test, laquelle a indiqué un état très émotif. On a évalué l'apprentissage (1) par la distribution des rats qui ont commencé à manger les semences, (2) par la distribution des rats qui ont atteint un apprentissage parfait, (3) par les erreurs (le nombre de semences non mangées), (4) le nombre moyen de semences mangées par chaque rat, (5) par le temps moyen mis par chaque rat pour manger une semence. (2) a été le meilleur pour la comparaison collective parce qu'il a représenté le rendement du groupe atteignant les critères de l'apprentissage. (1) a représenté l'apprentissage qui a atteint une partie des critères. (4) et (5) ont montré quelques phases de l'apprentissage. (3) a le plus différencié les individuels. L'habileté a été montrée par les résultats de temps (5) quand le temps moyen mis pour manger une semence est devenu moins que la limite de temps permise par les critères. L'habileté a été aussi montrée par le nombre d'enveloppes de semences mangées, lequel est devenu relativement constant après un certain jour.

On conclut qu'on peut différencier l'habileté de l'apprentissage dans une expérience d'apprentissage chez le rat quand les critères de l'apprentissage sont non équivoques ou quand on peut isoler quelques phases du rendement qui indiquent l'habileté, et que l'évaluation de l'apprentissage et la base de cette évaluation devraient être faites avec choix dans un certain but, c'est-à-dire, pour la comparaison collective ou pour la différenciation individuelle.

YOSHIOKA

GEWANDTHEIT UND LERNEN BEI RATTEN

(Referat)

Es wurde 33 Ratten, ungefähr 4 Monaten alt, täglich für 15 Tage die einfache Aufgabe gegeben in zehn Minuten zehn Sonnenblummensamen zu essen, in einer Umgebung, in der sie noch nie gegessen hatten. Der Grad der Gewohnung (habituation) an die neue Situation wurde gemessen (scored) an der Häufigkeit der Kotentleerung während der Prüfung. Diese Häufigkeit wies auf die Stärke der affektiven Aufregung hin. Das Lernen wurde gemessen (1) an der Verteilung (distribution) der Ratten die anfingen, die Samen zu essen, (2) an der Verteilung der Ratten, die vollkommenes Lernen erreichten, (3) an der Fehlerzahl (Zahl der ungefressen gelassenen Samen), (4) an der Durchschnittszahl der von jeder Ratte gegessenen Samen, (5) an der durchschnittlich zur Verzehrung eines Samens per Ratte verbrauchte Zeit. Der Masstab (2) erwies sich als der beste für Gruppenvergleichungen, weil er die Gruppenleistung darstellte, die die Kriterien des Lernens befriedigte. (1) stellte Lernen da, welches ein Teil der Kriterien befriedigte. (4) und (5) zeigten einige Phasen des Lernens. (3) differenzierte die Individuen am meisten. Gewandtheit wurde durch den Zeitverbrauch (5) erwiesen, wenn die durchschnittliche, zum Fressen eines einzigen Samens in Anspruch genommene Zeit geringer

wurde, als die ausserste, den Kriterien nach erlaubte Zeit (time limit). Gewandtheit wurde ebenfalls erwiesen an der Zahl der Hulsenfragmente der gefressen Samen, eine Zahl die nach einem gewissen Tag relativ konstant wurde. Man zieht den Schluss, dass die Gewandtheit von dem Lernen in einem Experiment mit Ratten unterschieden werden kann wenn die Kriterien des Lernens unzweideutig sind oder wenn einige Phasen der Leistung die auf Gewandtheit hinweisen isoliert werden konnen, und dass die Berechnung (scoring) des Lernens und die Basierung solcher Berechnung selektiv unter Besichtigung eines bestimmten Zieles, d.h., fur Gruppenvergleichung oder fur Differenzierung der Individuums, stattfinden soll

YOSHIOKA

AGE AND SEX DIFFERENCES IN THE TOY-PREFERENCES OF YOUNG CHILDREN*[1]

From the College of Education of the University of Minnesota

HAROLD BENJAMIN

In this study an attempt was made to examine the toy-preferences of young children, aged two to six years, by observation of their choice of toys in a standard situation Six toys were selected on the basis of similar size, cost, color,[2] and general attractiveness for subjects of the particular age group studied A description of these toys is given in Table 1.

TABLE 1
CERTAIN FEATURES OF SIX TOYS USED IN PREFERENCE SET-UP

Toy	Longest dimension	Next longest	Third longest	Movable parts	Chief color	Other colors	Cost
Car[3]	7 0 cm	3.5 cm	2 8 cm	wheels	carmine	gilt	10c
Girl[4]	8 0 cm.	2 5 cm	2 4 cm	arms	carmine	flesh yellow	10c
Horse[5]	7 8 cm	7 6 cm	2 6 cm	rider's arms	carmine	gilt	15c
Powder[6]	11.5 cm	4 0 cm	3 0 cm	cover	carmine	coral	15c
Airplane[7]	9 2 cm	7 5 cm	1 9 cm.	wheels, propeller	carmine	aluminum	10c
Boy[8]	8 0 cm	3 0 cm.	2 5 cm.	head	carmine	brown, white	10c 10c

The extent to which these toys were comparable may be further

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, February 11, 1932

[1]The writer is indebted to his former teacher and colleague, Lewis M Terman, for advice in connection with this study

[2]The toys were painted a standard color (waterproof carmine) by the writer

[3]A sedan automobile of composition metal

[4]A girl doll of bisque

[5]A mounted cowboy of composition metal

[6]A cylindrical vanity case of celluloid

[7]A monoplane of composition metal

[8]A boy doll of bisque

estimated by reference to the photographs repioduced in Figures 1 and 2

The following proceduie was used in presenting the toys·

1. Peimission to use a child as a subject was secuied by diiect request to one oi both of the parents The observer merely said that he was trying to find out what kind of toys children weie interested in and piomised to give a toy to each child who acted as a subject.

2. Childien were tested in their homes in the piesence of parents or other adults familiar to them Adults were requiied to maintain silence during the period of play All othei childien were excluded fiom the room.

3. The writer said to the child, "Would you like to look at the little playthings I have in this box? Hide your eyes as you do when you play hide-and-seek Turn youi back to me and hide youi eyes Don't peek. I am going to show you these little toys as soon as I get them ready." If necessary, the mother was asked to hold the child's head so he could not see the toys until they were set up

4. The toys were set up on a chair or a child's table so that they would be on a level with the child's bieast. They weie always presented as indicated in Figuie 1

5. The writer said, "All right Now you can look Play with any of them you want to" If the child hung back, the writei repeated, "They are fine to play with. Play with any one of them you like."

6. The amount of time the child played with each toy was recorded to the nearest second

7. At the end of five minutes, the observer took the toys from the child and again set them in the starting position The child was then placed in a position diiectly in fiont of the iow of toys at a distance approximately equal to the child's height. As the obseiver did this he said, "Now stand right there, please, so you can get a good look at all the toys. Which one do you like the best? I am going to give you the one you like the best" At this point the response was usually immediate. In the case of ceitain younger children, two and three years old, it was necessaiy to have the mother say, "Bring the one you like best to me. The man will give you one to keep, but he has to take the others away"

8 The observer secured auxiliaiy data by questioning the parent after the child had made a choice

The average time per subject for going through the above steps was approximately 30 minutes.

FIGURE 1
Six Toys Set for Presentation to the Subject

FIGURE 2
Another View of Six Toys Used in the Study

One hundred subjects were secured by selecting blocks at random in the city of Palo Alto, California, and vicinity, canvassing the houses in succession. As soon as the children available in an immediate neighborhood had been tested, the observer moved on to another neighborhood situated at a distance ordinarily too great for children of these ages to have become acquainted through neighborhood play (about 10 blocks). This precaution was taken to avoid the danger of having children discussing the situation with friends who had already received one of the toys. When possible, neighborhood groups of children were collected with the assistance of their mothers and were run through the test one by one in such a way that those who had already received toys were segregated until all were tested. Such precautions were, of course, always taken with children of one family.

Children of white parentage only were used. This fact constituted the only conscious basis of selection, aside from the age limits imposed. Various neighborhoods, ranging from the wealthiest to the humblest in the particular community, were used.[6]

No children were included who had ever attended any school other than the kindergarten. This fact set the upper age limit at approximately six years. The lower age limit was set by the necessity of getting a definite choice from the child. Attempts were made to get choices in the cases of several very young children (10 months, 11 months, and 12 months) by the method of taking one toy at a time away from the child and estimating the amount of displeasure shown. The method was found to be unworkable in the particular instances in which it was tried. Definite choice *was* shown, however, in the cases of the youngest children used in this study (14 months).

The age distribution of this group of children is given in Table 2.

Barr Scale ratings of the occupations of the fathers of the members of the group studied are given in Table 3. The mean Barr Scale rating of 10.30 may be compared with a mean rating of 7.92 estimated for the total adult male population of San Francisco, Oakland, and Los Angeles by Terman and Goodenough (1, p. 71). Since the United States Census does not furnish detailed occupational reports for communities of the size of Palo Alto, it was not possible

---

[6]Only one parent refused to allow a child to take the test, although many parents had to be assured that the writer was not a toy salesman with a new high-pressure selling scheme.

TABLE 2
AGE DISTRIBUTION OF ONE HUNDRED CHILDREN ACTING AS SUBJECTS IN THIS STUDY

Age in months	No of boys	No of girls	Total
14	1	0	1
15	1	1	2
16	1	0	1
17	0	1	1
18	0	1	1
19	0	1	1
20	2	0	2
22	0	1	1
23	1	0	1
25	0	2	2
26	1	2	3
27	0	1	1
29	2	0	2
30	1	0	1
31	2	3	5
33	1	0	1
34	0	1	1
36	1	2	3
37	0	3	3
38	4	1	5
39	1	3	4
40	2	0	2
42	0	2	2
43	1	1	2
45	1	3	4
46	0	2	2
47	2	0	2
48	1	1	2
49	5	0	5
50	1	0	1
51	1	1	2
52	1	3	4
53	0	2	2
55	1	1	2
56	2	1	3
57	1	1	2
58	2	1	3
59	1	1	2
60	1	2	3
61	4	1	5
62	1	0	1
67	0	1	1
68	1	1	2
69	2	0	2
72	0	1	1
76	0	1	1
Totals	50	50	100
Means	44 46 mos	43 78 mos	44 12 mos

TABLE 3

BARR SCALE RATINGS OF THE OCCUPATIONS OF FATHERS OF ONE HUNDRED CHILDREN

Rating	No of boys	No of girls	Total
17 81	4	1	5
16 71	0	2	2
15.75	1	2	3
15 71	1	1	2
15.14	1	2	3
15 05	3	0	3
13.71	1	1	2
13 31	1	1	2
13 29	1	0	1
13 21	2	1	3
11.74	2	4	6
11 71	1	1	2
11 51	3	2	5
11 34	1	4	5
11 17	4	0	4
10 26	2	2	4
9 72	1	1	2
9.37	4	6	10
8 99	1	2	3
8 75	1	0	1
8 08	1	1	2
7.91	0	1	1
7.79	0	1	1
7 73	0	1	1
7 54	0	1	1
7 24	0	1	1
6 27	5	3	8
5 81	3	4	7
4.98	4	1	5
3 99	0	1	1
3 62	2	2	4
Totals	50	50	100
Means	10 50	10 10	10 30

TABLE 4

PERCENTAGES OF FATHERS OF SUBJECTS OF THIS STUDY IN FOUR OCCUPATIONAL GROUPS COMPARED WITH PERCENTAGES OF ADULT MALE POPULATION OF TWO CITIES IN SAME GROUPS

Occupational group	Present study	San Francisco and Los Angeles*
Professional	17	2 9
Commercial	27	36 1
Industrial	54	57 7
Public Service	2	3 3

*Figures in this column are taken from L M Terman *et al* (1, p 63)

to estimate whether the children in this group were far above the social level of the average in the community. It is probably fair to assume that they were somewhat above that level A comparison of the percentages of fathers falling in the four United States Census groups with percentages of adult males in San Francisco and Los Angeles in the same groups shows a decided overemphasis on the professional group among the subjects of this study, partly traceable to a university population in Palo Alto and neighboring towns

All the children in this group were born in the United States, with one exception—a child born in England The racial composition of the group is indicated in Table 5

Tables 6 and 7 show the final selections made by the one hundred children after playing with the toys for five minutes.

Determining the significance of the differences between the choices of the two sexes, we secure the results presented in Table 8

TABLE 5
BIRTHPLACES OF FATHERS OF SUBJECTS OF THIS STUDY

Country	Boys	Girls	Total
United States	40	40	80
Ireland	3	5	8
Austria	3	0	3
France	2	1	3
England	1	1	2
Germany	1	1	2
Sweden	0	2	2
Totals	50	50	100

TABLE 6
FINAL TOY CHOICES OF FIFTY BOYS

Ages in months	Car	Airplane	Horse	Powder	Boy	Girl	Totals
62-69			1	2	1		4
54-61	3	6	3				12
46-53	4	6		1			11
38-45	4	3	1	1			9
30-37	2	1	1	1		1	6
22-29	2			1			3
14-21	2	1			2		5
Totals	17	17	6	6	3	1	50

TABLE 7
FINAL TOY CHOICES OF FIFTY GIRLS

Ages in months	Car	Airplane	Horse	Powder	Boy	Girl	Totals
70-76		2					2
62-69		1				1	2
54-61			1	2	1	4	8
46-53	3	4	1	1			9
38-45	1	3		2	2	2	10
30-37	2	1	1		1	4	9
22-29	1	1			1	3	6
14-21					3	1	4
Totals	7	12	3	5	8	15	50

TABLE 8*
SIGNIFICANCE OF DIFFERENCES BETWEEN TOY-PREFERENCES OF BOYS AND GIRLS

Choice	Proportion of boys choosing	Proportion of girls choosing	Diff	σ_d	$\frac{Diff.}{\sigma_d}$	Chances in 1000 that difference is significant
Car	34	14	.20	083	2 421	992
Airplane	.34	.24	10	090	1 111	867
Horse	.12	06	06	057	1 053	854
Powder	.12	10	02	063	317	624
Boy	06	16	10	.062	1 613	947
Girl	02	30	28	068	4 118	999 plus

*The values given in the fifth column were found by use of the formula $\sigma_{(p-p^1)} = \sqrt{\sigma_p^{\,2} + \sigma_{p^1}^{\,2}}$ in which $\sigma_p^{\,2} = \frac{pq}{N}$ and $\sigma_{p^1}^{\,2} = \frac{p^1q^1}{N}$ The chances in 1000 that the differences were significant were secured by entering the Kelley-Wood Table of the Normal Probability Integral with the values given in the $\frac{Diff.}{\sigma_d}$ column

The results shown in this table clearly indicate that *girl* is distinctly a feminine choice with our group of children Only one boy selected this doll *Boy* was also fairly feminine as a choice The third toy designed to catch the feminine eye, the vanity case, was a failure in this respect The difference between the sexes in the choice of this toy is only a little more significant than a fifty-fifty chance, and what little significance there may be is in favor of the boys.

Car comes nearest to being the most distinctly masculine toy in

this group, with a difference of 2 421 sigmas Although the *airplane* is probably significantly masculine, it has a strong feminine following well distributed through the various age groups. *Horse* is almost as significantly masculine

Referring to the column totals in Tables 6 and 7, further details may be noted concerning the relative attractiveness of the toys. Almost 70% of the boys divided their choices equally between the two transportation toys, *car* and *airplane*. A substantial group divided their choices again between the ornamental toys, while an almost negligible group were concerned with the dolls. Of the four boys who chose dolls, it should be noted, two were 20 months old and one was 29 months old. The fourth boy was 69 months old, however, and the age distribution of doll choice may, of course, be due to chance alone

The strong interest of the girls in *airplane* is responsible for rais-

TABLE 9*
TOYS WITH WHICH BOYS PLAYED LONGEST

	Car	Airplane	Horse	Powder	Boy	Girl	Totals
Played longest with final choice	10	5	2	2	1		20
Played longer with one toy than with final choice	2	2	2	2			8
Played longer with two toys than with final choice	4	3	3		4	2	8
Played longer with three toys than with final choice	5	4	5	5	4	1	8
Played longer with four toys than with final choice	2	2	2	2	2	2	3
Played longer with five toys than with final choice	2	3	3	2	2	3	3

*The numbers in the columns headed *Car*, *Airplane*, etc, refer to the number of times the toys so designated was given more time than was given to the child's final choice The figures in the column headed *Totals* refer to the number of children who played longest with final choice, longer with one toy than with final choice, longer with two toys than with final choice, three toys, etc

TABLE 10

TOYS WITH WHICH GIRLS PLAYED LONGEST

	Car	Airplane	Horse	Powder	Boy	Girl	Totals
Played longest with final choice	3	5	1	1	5	5	20
Played longer with one toy than with final choice	4	3	1		4	2	14
Played longer with two toys than with final choice	2	3	1	2	1	1	5
Played longer with three toys than with final choice	4	3	4	4	4	2	7
Played longer with four toys than with final choice	1	2	1	2	2		2
Played longer with five toys than with final choice	2	1	2	2	2	1	2

TABLE 11

SIGNIFICANCE OF DIFFERENCES BETWEEN AMOUNT OF TIME BOYS AND GIRLS PLAYED WITH EACH TOY

Toy played with longest	Proportion of boys	Proportion of girls	Diff	σ_d	$\frac{Diff}{\sigma_d}$	Chances in 1000 that difference is significant
Car	34	.20	14	087	1.609	946
Airplane	22	.22	00	—	000	500
Horse	14	08	06	063	952	829
Powder	18	08	10	055	1 818	965
Boy	10	26	16	075	2 133	983
Girl	02	16	.14	056	2 500	994

TABLE 12

SIX TOYS RANKED ACCORDING TO THEIR PLAY INTEREST FOR BOYS AND GIRLS

Toy	Interest for boys	Interest for girls
Car	1 (69 82 seconds)	4 (46.10 seconds)
Airplane	2 (58 30)	3 (46 76)
Horse	3 (52 96)	6 (32 08)
Boy	4 (39 70)	1 (51.82)
Powder	5 (34 32)	5 (37 18)
Girl	6 (23 52)	2 (50 20)

ing the total choices for that toy well above its nearest competitor. Although *girl* leads in number of feminine choices, it is a question whether *airplane* has not as much feminine interest.

The length of time the children played with the various toys during the course of each presentation furnishes a second index of their interests. It might seem plausible to assume that a child would usually select the toy with which he played longest. As a matter of fact, 60% of the children played longer with one or more toys other than the one finally selected. Tables 9 and 10 show what toys were chosen as play objects during the five-minute period but were not taken in the final selection. From a consideration of the data presented in these tables, we see that *car* and *horse* remain definitely masculine in their appeal, while *powder* changes from the neutral ground shown in Table 8 to masculine territory. *Boy* and *girl* remain feminine. *Airplane* appeals equally to both sexes.

The significance of differences between sexes in amount of time played with each toy is shown in Table 11.

Ranking the toys according to attractiveness for boys and girls, as measured by the mean number of seconds spent in play with each toy, we have the results given in Table 12.

In an attempt to discover a possible relationship between age and the degree of masculinity of interest as represented by the ranking of toys according to the length of time which boys devoted to them during the play period, correlations were worked out between age in months and ranking in the first column of Table 12 for boys and girls. In both sexes a small negative correlation was found; for boys, —.26, for girls, —.22. Perhaps this indicates a slight tendency for masculinity of interest to increase with age from one year to six.

No significant relationship could be found between number of sibs and choice of toys or between age and "decidedness" of choice.[10] Correlations between age and number of times the child changed from one toy to another during the period of play were insignificant, —.08±.15 for the girls and —.21±.14 for the boys.

In 72 cases the observer was able to secure a statement from the child's mother as to whether she would have foreseen the child's choice. It was impossible to secure this statement in all cases, since the observer had to bring the matter up in casual conversation at the conclusion of the test. In every case the mother appeared anxious

[10]"Decidedness" was defined as tendency to play with one toy longest and select it as the final choice.

to show that she knew her child's likes and dislikes very well, yet 29 mothers testified that they were astonished at the child's choice. The mother of the 30-months-old boy who selected the distinctively feminine toy *girl* appeared to be deeply chagrined at the child's evidence of feminine interest. She assured the writer that the boy had never before given her any cause for serious anxiety.[11]

Summary

1 Three toys used in this study show conclusive sex differences, whether measured by final choice or by length of time with which children handled them during the play period: *car*, a masculine toy; and *girl* and *boy*, feminine toys The three remaining toys, *airplane*, *horse*, and *powder* are of little value in showing sex difference

2 Sex differences in choice of *car*, *girl*, and *boy* were found throughout the various age levels of the group studied

3 In 60% of the cases it was impossible to predict the final choice from the amount of time spent in playing with particular toys

Reference

1 Terman, L M, *et al* Genetic studies of genius' Vol I Mental and physical traits of a thousand gifted children Stanford University, Calif Stanford Univ Press, 1925 Pp xv+648.

University of Minnesota
Minneapolis, Minnesota

LES DIFFÉRENCES D'ÂGE ET DE SEXE DANS LES PRÉFÉRENCES DES JEUNES ENFANTS POUR LES JOUETS

(Résume)

Cet article rapporte les préférences de cent enfants, âgés de quatorze à soixante-seize mois, pour six jouets, une boite à poudre, une poupee en forme de garçon, une poupée en forme de fillette, une automobile, un avion, et un cavalier, appelés respectivement, Poudre, garçon, fillette, voiture, avion, et cheval Trois d'entre ces jouets ont montré des différences conclusives de sexe, mesurées ou par le choix final ou par la duree où

[11]This boy played longest with *car* As he ran the toy back and forth along the floor, however, he appeared to be looking in the direction of *horse* and *girl* Then he played with *horse* in the same manner, always with an eye on the remaining toys On one occasion he picked up all the toys with the exception of *powder*, and after 42 seconds released them to play with *horse* alone In this, as in many other cases, it is hard to find any evidence in the child's play which would indicate what his final choice would be. Yet his decision was made with surprising promptness at the end of the play period.

les enfants s'en sont servis pendant une période étalonnee de jeu de cinq minutes, voiture, jouet masculin, et fillette et garçon, jouets féminins Les trois autres jouets ont montre très peu de differences de sexe On a trouve des différences de sexe dans le choix de voiture, fillette et garçon à tous les divers niveaux d'âge du groupe d'enfants étudié. Il n'a été possible de prédire le choix final de l'enfant selon le temps passé à jouer des jouets specifiques que dans environ quarante pour cent des cas On n'a découvert aucune relation signifiante entre l'âge et le degré de masculinité de l'intérêt Il est très probable que les parents des enfants des âges étudiés dans cette etude, tendent ordinairement a surestimer leurs connaissances des préférences et des antipathies des enfants.

BENJAMIN

ALTERS UND GESCHLECHTSUNTERSCHIEDE IN DEN SPEILZEUG-BEVORZUGEN KLEINER KINDER

(Referat)

In dieser Schrift wird Bericht erstattet uber die Bevorzugungen von 100 Kindern 14 bis 76 Monate alt, in Bezug auf sechs Spielzeuge eine Puderschachtel (vanity case), eine Knabenpuppe, eine Madchenpuppe, ein Automobil, ein Luftschiff, und einen Reitersmann, respektiv Puder, Knabe, Madchen, Auto (car), und Pferd benannt An drei dieser Spielzeuge liesen sich bestimmte Geschlechtsunterschiede erweisen, die sowohl an der endgultigen Wahl wie an der Zeit, wahrend der die Kinder sie innerhalb einer Normalspielzeit von 5 Minuten behandelten, erwiesen werden konnten, namlich das *Auto*, ein Knabespielzeug, und der *Knabe* und das *Madchen*, Madchenspielzeuge. Der drei ubrigen Spielzeuge waren bei der Demonstrierung von Geschlechtsunterschieden von kleinem Wert Geschlechtsunterschiede in Bezug auf die Wahl des Autos, des Madchens, und des Knabens zeigten sich durch alle Altersniveaus der untersuchten Kindergruppe hindurch In nur 40% der Falle konnte man die endgultige Wahl des Kindes voraussagen aus der Zeit die es an besonderen Spielzeugen verwendete. Man entdeckte keinen bedeutenden Zusammenhang zwischen dem Alter des Kindes und dem Grad der Mannlichkeit (masculinity) der Interessen Es ist hochst wahrscheinlich, dass Eltern von Kindern in der in dieser Untersuchung beobachteten Altersgruppe geneigt sind, den Umfang ihres Wissens um die Vorlieben und Abneigungen der Kinder zu uberschatzen.

BENJAMIN

FOOD PREFERENCES OF THE ALBINO RAT*

From the Psychological Laboratories of the University of Wisconsin

H F Harlow[1]

There is a considerable amount of evidence indicating that the mechanism of appetition functions in a specific manner to bring about the ingestion of food substances necessary for survival It is well known that animals, as well as human beings, manifest "hungers" for particular kinds of food, but most of the data available are of an observational nature The present experiment was planned to determine the degree of preference shown by albino rats for various kinds of food, and to attempt to discover some of the causal factors operative in the choice.

I Pertinent Literature

Bayer (1) carried out an extensive study of the conditions influencing ingestion in the chicken. For the determination of food preferences he made use of two methods which he calls, respectively, "direct" and "indirect." In the direct method only one kind of food was presented and the animal was allowed to feed to satiation The procedure was carried out three times with each of the foods tested: rice, rye, barley, wheat, and oats Inasmuch as the nutritive value of the food might enter to influence the quantity ingested, he then contrived the indirect method Five grains of each of two kinds of cereals were mixed and placed before the animal, and the first grain chosen was observed Since the total amount of food ingested by this method was small, factors of nutrition and decreasing appetite were probably kept at a minimum. The two hens tested carefully showed consistent preference of rice to rye, barley, wheat, or oats, and four more animals confirmed this finding Some temporal variation in preference occurred, but the most and least preferred grains showed least fluctuation in the case of the individual animal

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, March 10, 1932

[1]These experiments were carried out by the author at Stanford University under the direction of Dr C P Stone. The work was financed by the Thomas Welton Stanford Fund

Bayer also found that the animals tended to overeat when given rice for a three-day period; when tested for preferences by the indirect method following this period, they showed less preference for rice. Wheat did not suffer much when this procedure was tried, that is, presumably the animals did not overeat.

The investigations of Evvard (4), Osborne and Mendel (7), and Mitchell and Mendel (6) confirm the observational data on the ability of animals to choose diets containing adequate nutritive substances In general, the results of these investigators indicate that the animals chose from both adequate and inadequate diets and that most of the animals chose enough of the adequate food to insure an almost normal growth rate, but some animals clung to the inadequate diet. Probably the scarcity of animals choosing inadequate diets is due to the fact that they are rapidly eliminated under natural conditions

Simmons (8) found that bread and milk as a reward for maze running made for much greater efficiency of performance than sunflower seed, the difference between the two groups being 60% in errors and 94% in time

Careful observational reports of food preferences in apes have been reported by Yerkes (9, pp 212-214) and Zeigler (10, pp. 175-176) They indicate that these animals are quite similar to humans in their likes and dislikes of particular foods

II. Technique

The technique of the present studies was devised to follow that of Bayer (1) in his "indirect" method. Modifications were, of course, made to suit the testing of rats, and certain controls were instituted.

Male rats were weaned at 24 days of age and placed in groups of four or five in cages 18x18x8 inches No food was kept in the cages, but water was always available. Experimentation began on the following day, and continued for 20 consecutive days

During the experimental period the rats were put individually, twice a day, into the empty food cage (size the same as the home cage) and given 15 minutes to explore and accustom themselves to the new situation. Four small pans, each containing one kind of food, were then quietly introduced into the corners of the cage and the rat was allowed to feed undisturbed for an hour The position of the foods was altered each day to prevent preferential response to a certain position

Two criteria of food preference were used: first, the food which was first eaten in a given experimental period; and, secondly, the amount of each food actually consumed The foods used were: in Experiment I, standard diet (McCollom's diet), whole oats, whole corn, and sunflower seed, in Experiment II, standard diet, meat scraps, whole corn, and cheese; in Experiment III, standard diet, raw beef scraps, and whole corn; and in Experiment IV, the same as in Experiments I and II.

Male albino rats born in the laboratory were used in these experiments In Experiments I and II groups of 10 normal rats were used. in Experiment III, 9 intact rats, 11 blind animals, 8 anosmics, and 6 blind-anosmics; in Experiment IV, one group of 8 and one of 6 normal rats.

III. Results

Experiment I. In this experiment the rats, who had had no previous experience with solid food, were allowed to choose between standard diet, whole oats, whole corn and sunflower seed. According to both criteria of preference (first choice and total amount eaten), the order of preference is: standard diet, corn, oats, and sunflower seed

For the 20 days of the experiment the total number of first choices of the 10 rats were as follows: standard diet, 274, corn, 70, oats, 56, sunflower seed, 0. (The sum is 400, since there were two daily feedings.)

The average amounts eaten and the variability are shown in Table 1 The comparisons between means and the reliabilities of the differences are shown in Table 2.

The critical ratios showing the superiority of standard diet over any of the other foods are very large and indicate that the chance that this difference is not a true one is extremely small

Corn is preferred to oats, but the differences are quite unreliable None of the critical ratios reach the value of 1.00 Corn is preferred to sunflower seed at the beginning of the experiment, but the reliability of the difference diminishes toward the end Oats are also preferred over sunflower seed, the differences being reliable in two of four comparisons

Experiment II. This experiment was identical with the first, except that oats and sunflower seed were replaced by meat scraps (table scraps) and Kraft's limburger cheese.

TABLE 1
MEAN VALUES AND VARIABILITY IN AMOUNT OF VARIOUS FOODS CHOSEN

Food eaten	Age of rats in days 25-29	30-34	35-39	40-44
Standard diet:				
Mean	28.30	35 00	40 80	50 80
$S D_m$	1 43	1 26	1 87	2 95
Corn:				
Mean	1 40	4 40	4 00	2 10
$S D_m$	0 39	2 00	1 37	1 50
Oats				
Mean	0 80	2 75	2 50	0 85
$S D_m$	0 08	1 05	0 89	0 13
Sunflower seed:				
Means	0 00	0 20	0 02	0 00
$S D_m$	0 00	0 04	0 03	0 00

The order of preference in this experiment was standard diet, meat scraps, cheese, and corn The differences between standard diet and meat scraps were not completely reliable, but were consistent throughout the four five-day periods of the experiment The superiority of either of these foods over cheese and corn, and of cheese over corn, is statistically significant, the critical ratios all being larger than 3 00.

Experiment III Four sensory modalities may be involved in the discrimination shown by the preceding results. visual, olfactory, gustatory, and tactual The possibility of tactual discrimination seems quite unlikely. Since the taste receptors cannot conveniently be either inactivated or destroyed, it was decided to eliminate the possibility of vision and smell in an attempt to get at the sensory controls for this food preference.

All the animals used in this experiment were carefully prevented from having any experience with solid food before the beginning of the experiment They were weaned at the age of 21 days, and were then fed on a diet of diluted powered milk and tomato juice Under this regime the intact and blind rats weighed at 40 days about 10% less than the average, and the anosmic and blind-anosmic animals about 20% less The blind and anosmic groups were operated on

TABLE 2
RELIABILITY OF DIFFERENCES BETWEEN VARIOUS FOODS CHOSEN

Comparison	25-29	Age of rats in days 30-34	35-39	40-44
St diet—corn				
Diff.	26 9	30 6	35 2	48 7
SD_{diff}	1 48	2.36	2 31	3 31
C R	18 18	12 97	15 93	14 71
St diet—oats				
Diff.	27 5	32 25	38 3	49.95
$S.D_{diff}$	1 43	1 64	2.07	2 95
C R	19.23	19 66	13 67	16 93
St. diet—sunflower seed				
Diff	28 3	34 8	40 78	50 8
SD_{diff}	1 43	1 26	1 87	2 95
C R	19 79	26 62	21 71	17 22
Corn—oats.				
Diff	0 6	1 65	1 5	1 25
SD_{diff}	0 40	2.26	1 63	1.51
C R.	0 15	0.73	0 87	0 83
Corn—sunflower seed:				
Diff	1 4	4 2	3 98	2 1
$S.D_{diff}$	0 39	2 0	1 37	1 50
C.R	3 59	2 10	2 77	1 40
Oats—sunflower seed.				
Diff.	0 8	2 55	2 48	0 85
SD_{diff}	0 08	1 05	0 89	0 13
C R.	10 00	2 43	2 58	6 54

at 30 days of age, the blind-anosmic group were blinded at 26 days, and the olfactory bulbs were destroyed at 30 days.

The mean number of grams of food of each kind (meat, standard diet, and corn) taken by each group during the 10 days of the experiment is shown in Table 3 It will be noted that no characteristic differences exist between the intact, blind, and anosmic groups. The meat is slightly preferred to the standard diet by all three groups, and both of these foods are decidedly superior to the corn. The blind-anosmic group shows a slight preference for standard diet over meat, and is consistent in rejecting the corn

TABLE 3
FOOD CHOICES FOR OPERATED ANIMALS

Group	Meat	Kind of food Standard diet	Corn
Blind	58 2	53 1	11 2
Anosmic	63 4	54 0	11 4
Blind-anosmic	57 8	60 0	12 1

These results seem definitely to indicate that food preference as shown in the present experiment is not dependent upon either visual or olfactory cues It is, of course, possible that these facilitate the choice in the intact animal, but they clearly are not indispensable

Experiment IV The object of this experiment was to determine the effect of long experience with a single foodstuff upon the food preference of the rat

The animals were weaned at 21 days. Eight rats were then fed for 68 days on standard diet supplemented with a little lettuce Six rats were fed for 53 days on a diet of corn, lettuce, and whole milk, the milk being added to prevent a serious protein deficiency. In both cases the animals were given no food for one day before the first observation on food preference

Of the eight standard-diet animals, four were allowed to choose from standard diet, corn, oats, and sunflower seed and four from standard diet, meat, cheese, and corn In the first case, standard diet was preferred almost exclusively, the differences being even larger than in the case of Experiment I In the other choice, the animals ate over twice as much standard diet as meat—a surprising result when it is recalled that in Experiment II the two foods were approximately on a par The corn was almost completely ignored by this group.

Bayer (1) reports no case in which prolonged experience with a single food accentuated the appetite of his hens for this substance—the situation we have found in this experiment Bayer's hens, however, had had experience with other foods before the beginning of the experiment, so the results are not strictly comparable.

The six rats who had been fed on a low-protein diet (corn, milk, lettuce) for 53 days were given a choice of standard diet, meat, corn, and oats The results here are strikingly different from those of the standard-diet group Almost three times as much meat as standard diet was consumed by this group The total grams eaten

by the six rats in 20 days of each food were as follows: meat, 1162, *standard diet*, 486, *corn*, 189; *oats*, 85.

These results make it appear that the rats were the victims of a protein insufficiency during the corn-feeding period. This conclusion is supported by the fact that the animals at the beginning of the experimental period were more than 20% underweight The assumption of a specific protein hunger would explain the meat preference, as meat is much higher in protein than is the standard diet.

IV Discussion

These experiments show clearly that the albino rat discriminates *between different food substances and manifests definite food* preferences Standard diet and meat are much more frequently chosen by the animals than cheese, corn, oats, or sunflower seed. Sunflower seed is clearly discriminated against by these rats. Under conditions of protein insufficiency the animals choose a much higher protein diet than when they have been receiving a balanced ration

Like the results of a previous study on social facilitation of the feeding response of the white rat (5), these data open a question as to the nature of the factors determining ingestion The classical experiments of Cannon and Washburn (2) and Carlson and his students (3) have clearly demonstrated the rôle of the contractions of the stomach wall in producing the sensations associated with hunger. From this it might naively be assumed that stomach contractions are the only determining factor in hunger, although Cannon points out that "appetite" plays an accessory rôle

Our previous investigation of the rôle of social facilitation in increasing the amount of food ingested by white rats, and the present results showing clearly a discrimination and preference for foods which have a specific dietary value, make it clear that hunger and ingestion are not related in a simple one-to-one fashion The feeding response of the rat is a complex function and a large number of contributory factors must be analyzed before it is thoroughly understood.

V Summary

1. The male albino rat without previous experience with solid food still shows definite food preferences

2. These preferences are such as to make for a complete balanced diet (meat, standard diet) as opposed to an insufficient diet (corn, oats, sunflower seed).

3. Destruction of vision, of olfaction, and of vision and olfaction does not alter the food choices of the rat These are therefore probably based on gustatory (or tactual) cues.

4 Rats who have received an insufficient ration show a very marked preference for foods containing the elements of which they have been deprived. Rats who have been on an adequate diet for a long time may prefer this diet to another adequate diet.

References

1 BAYER, E Beitrage zur Zweikomponententheorie des Hungers *Zsch f Psychol*, 1929, 112, 1-54

2 CANNON, W B. Bodily changes in pain, hunger, fear and rage (Rev ed) New York Appleton, 1929. Pp 404

3. CARLSON, A J. The control of hunger in health and disease Chicago. Univ Chicago Press, 1916 Pp. viii+319

4 EVVARD, J M Is the appetite of swine a reliable indication of physiological needs? *Proc Iowa Acad Sci.*, 1915, 22, 375-403

5 HARLOW, H F Social facilitation of feeding in the albino rat *J. Genet Psychol*, 1932, 41, 211-221

6 MITCHELL, H S, & MENDEL, L B The choice between adequate and inadequate diet, as made by rats and mice *Amer J Physiol*, 1921, 58, 211-225

7 OSBORNE, T B, & MENDEL, L B The choice between adequate and inadequate diets as made by rats *J Biol Chem*, 1918, 85, 19-27

8 SIMMONS, R The relative effectiveness of certain incentives in animal learning *Comp Psychol Monog*, 1924, 2, No 7 Pp 79

9. YERKES, R. M Almost human New York Century, 1924 Pp. xxi+278

10. ZEIGLER, H E Der Begriff des Instinktes Einst und Jetzt Jena Fischer, 1920.

University of Wisconsin
Madison, Wisconsin

LES PRÉFERENCES DE NOURRITURE DU RAT BLANC

(Résumé)

Les rats blancs récemment sevrés montrent des préférences de nourriture lesquelles restent tout-à-fait constantes pendant une période de 20 jours On peut montrer cela en mesurant ou la quantité de nourriture mangée pendant un temps donne ou le nombre de premiers choix pour les aliments individuels. Les deux mesures donnent le même ordre de préférence le régime ordinaire (régime de McCollum) est preférable au ble d'Inde, le ble d'Inde à l'avoine, et l'avoine aux semences de tournesol De même, le régime ordinaire est préférable à la viande, la viande au fromage, et le fromage au blé d'Inde La plupart des différences sont statistiquement

constantes (CC de 3 ou plus) sauf celles entre le régime ordinaire et la viande et entre l'avoine et le blé

La préférence du régime ordinaire et des morceaux de viande au blé peut se montrer chez le rat blanc même après la destruction des récepteurs visuels, des bulbes olfactifs, et celle de tous deux de ces organes de sens. Cela le rend probable que le choix peut être fait sur la base de repères gustatifs et tactiles

La préférence pour un régime riche en protéines semble être accentuée dans un groupe de 14 rats qui ont un régime peu riche en protéines pendant 53 à 68 jours Cela indique que les préférences de nourriture peuvent être influencées par les besoins corporels changeants des animaux

HARLOW

DIE FUTTERBEVORZUGUNGEN DER ALBINORATTE

(Referat)

Frisch entwohnte Albinoratten offenbaren Futterbevorzugungen, die wahrend 20 Tagen ziemlich konstant bleiben Man kann diese Bevorzugungen demonstrieren, indem man entweder die Quantitat des innerhalb einer festgesetzten Zeit verzehrten Futters, misst, oder bestimmt, wie oft jedes Futter zuerst gewahlt wird. Man erhalt mit den beiden Verfahren die selbe Rangordnung der Bevorzugungen, die Normalkost (Kost von Mc-Collum) wird dem Mais vorgezogen, der Mais dem Hafer, und der Hafer dem Sonnenblumensamen Anlich wird die Normalkost dem Fleisch vorgezogen, das Fleisch der Kase, und die Kase dem Mais Die Mehrzahl der Unterschiede sind statistisch zuverlassig (Koeffizient der Zuverlassigkeit' 3 oder mehr), mit Ausnahme der Unterschiede zwischen Normalkost und Fleisch und zwischen Hafer und Mais

Die Bevorzugung von Normalkost und Fleischstuckchen uber Mais kann in der Albinoratte bewiesen werden sogar nachdem die visuellen Rezeptoren oder die Geruchswurzeln (olfactory bulbs) oder beide dieser Sinnesmodalitaten zerstort worden sind Es scheint also wahrscheinlich zu sein, dass die Wahl auf Geschmack- oder Tastanweisungen basiert ist

Die Bevorzugung einer proteinreichen Kost scheint verstarkt zu sein in einer Gruppe von 14 Ratten, die wahrend 53 bis 68 Tagen mit proteinarmer Kost gefuttert worden sind Dieser Befund weist darauf hin, dass die Futterbevorzugungen durch wechselnden korperlichen Bedurfnisse der Tiere beeinflusst werden konnen

HARLOW

A CLINICAL STUDY OF THE APPLICATION OF MENTAL HYGIENE TO THE TREATMENT OF CHILDREN'S SCHOOL PROBLEMS*

From the Department of Psychiatry of Columbia University

JOHN LEVY

At the present time when the value of a mental hygiene approach to the treatment of various types of maladaptation is being more critically considered than heretofore, a study which attempts to measure concretely the value of this approach in the field of scholastic maladjustment would seem to be of timely interest.

The following investigation of the question of what mental hygiene can do to eliminate the friction existing between school and scholar was carried out by means of an intensive clinical study of the problems of 36 children whose ages varied from 9 to 17 years Twenty cases were boys The children were selected in equal numbers from public school and high school Almost all of the parents of the children studied belonged to upper and lower middle-class economic groups. The occupations of these parents varied from salesmen and mechanics to dentists, policemen, and cooks

The intelligence of the children was above the average, the median IQ being 112 5 There were only 9 cases with IQ's of 105 or lower, and only one case below 90. This was a border-line defective girl with an inadequate personality who was referred because of shyness and antagonism toward authority. Two cases rated around 140. One of them was a young boy of 9 who was referred because of stealing, lying, and personality difficulties. The other case was a boy of 12 in junior high school whose teachers thought he was unhappy and needed help While, of course, one would expect that high-school clientele would turn in a good many high IQ's for study, it must also be kept in mind that a good many children with average intelligence attend at least the earlier years of high school That children of high intelligence are having difficulties in school again calls attention to the fact that superior intelligence is no safeguard against school maladjustment.

*Accepted for publication by Carl Murchison of the Editorial Board and received in the Editorial Office, January 20, 1932

TABLE 1
PROBLEMS AS DIAGNOSED BY REFERRING SCHOOL

Interference with school routine	23 cases
Poor adjustment with other children	12 cases
Poor school work	11 cases
Dishonesty	10 cases
Sex difficulties	2 cases

A classification (Table 1) of the various problems as diagnosed by the referring school before they were studied by the clinic throws light on the outward form of children's behavior which schools regard as worthy of clinical attention. It throws light, too, upon the type of conformity school authorities are anxious to obtain.

As is to be expected, the school is more interested in children who do not "fit in" than in any other type of problem child. Almost twice as many children were referred because of their inability to keep in step than because of poor school work. It is encouraging, though, to see that schools are also interested in children's adjustment to other children. Only two children were referred because of sex problems. Perhaps children are getting away with a good deal of freedom which is not being discovered.

Coming under the heading of "interference with school routine" were such problems as truancy, difficulties of one kind or another with teachers, restlessness, irritability and bids for attention. One child interfered with the routine because she was "too nervous to recite in class." Incidentally, this girl turned out to be a case of post-traumatic psychosis. One boy interfered with the routine in class by his "tendency to spend much time in sleep." This boy was suffering from the effects of a previous attack of infantile paralysis. The more usual type of interference with school routine is seen in the case of a boy who was referred because "he was easily irritated, becomes very violent when angry. He is nervous, jerky and fidgety."

Those cases referred because of their poor academic adjustment included children with specific disabilities, e.g., poor mathematical ability, grade placement disproportionate to their intelligence, or other types of scholastic backsliding. One girl was failing in "four out of five of her major subjects." A boy did not want to complete his high-school course "because of his ardent desire to go on the stage." One girl had failed in Spanish on successive occasions even though her IQ was above 110. This girl had a bad emotional tie-up with her

language teacher, which blocked her in her work "He is not doing satisfactory work and is inattentive" were frequent causes for referring children to the clinic

The reasons given by teachers for referring to the clinic children classified under "poor adjustment with other children" were "no friends," "unhappy," "change in personality," "cruel," and "timidity." One girl was studied because "she had frequent spells of crying" Another girl was referred because "she hits her brother." A high-school boy was sent in "because he keeps to himself." This boy was suffering from dementia praecox E H suffered from an "entire lack of social adaptation He spends his time travelling around on Brooklyn street cars all day, even though he is 14 years of age and old enough to know better" E C. was referred because she had only one girl friend, and an over-intimate relationship with this girl was suspected W N thought himself "too superior to mix with anybody and told the teachers that he thought he knew more than they did" E. B, a girl of 12, insisted on playing only with much older girls Her parents thought that she was dangerously on the verge of sex delinquency J D. insisted upon hobnobbing only with actors and men interested in the stage. His school adviser suspected homosexual tendencies, an opinion which was confirmed by further clinical study.

The problems of dishonesty included not only serious lying but frank stealing One girl was referred because "she had stolen nine pocketbooks out of a schoolroom" A boy of 15 was referred for "forgery of checks totalling $28 in order to carry on truancy." E C., a girl of 12, "steals from her home and recently was involved in a stealing episode in school" Another girl was "caught in the act of stealing 50c from her father's pants pocket the first week he was married to the stepmother This term she stole $10 from her teacher's pocketbook"

While it might have been interesting to study the type of problem that children of different ages present, and the sex differences in problems presented, such a study is best done on a large quantitative basis and is precluded at the present time by the limited number of cases in the present study

It is interesting to compare with the labels that the school authorities fasten upon the behavior deviations of children the psychiatric diagnoses obtained after adequate histories on the children's behavior have been studied and full examinations (physical, psychological, and

TABLE 2
PSYCHIATRIC DIAGNOSIS

Psychoses	5 cases
Neuroses	1 case
Personality disturbances	12 cases
Domestic disturbances	7 cases
Feelings of insecurity	4 cases
Sex conflicts	2 cases
Developmental anomalies	2 cases
Sociological problems	2 cases
Adolescent revolt	1 case

psychiatric) have been carried out. Table 2 gives the psychiatric diagnoses for problems discussed above.

It must be kept in mind that a diagnosis is only a restatement from a different point of view of the problem which the child presented; it is a convenient psychiatric label used as a means of conveying a certain connotation. These diagnoses do not state the underlying causes of the behavior; such causes will be discussed subequently. Moreover, it is quite possible that psychiatrists would not altogether agree with this choice of stickers, except perhaps in the matter of the psychoses.

These psychoses included one case of dementia praecox, another case of manic depressive insanity which had been previously diagnosed schizophrenia, one case of encephalitis complicating infantile paralysis. The other psychoses were post-traumatic. The neurosis diagnosed was the case of a young boy, 8 years old, whose nervous system had been shot to pieces by emotional wear and tear in the home.

The sex conflicts were based upon latent homosexuality. The adolescent revolt was a case of a boy of 15, well-developed physically, who was trying to cut his mother's apron strings against her protest. Sociological problems included impossible home and neighborhood conditions due to the nature of their social and economic make-up. Two boys from very questionable stock whose birth and early development were highly traumatic were classified as developmental anomalies.

One of the cases classified under "feeling of insecurity" was a young girl, aged 10, who had been brought up amid terrific marital discord. She had been a witness to repeated attempts of her father to murder her mother. The "feeling of insecurity" in another case was brought about by the patient's mother who was vying with the patient for

her brother's attention L. B's feeling of insecurity was due to the fact that he was an orphan and not wanted in his relative's home

Personality inadequacies varied from the hysterical personality which J D expressed in fainting attacks at school to the shyness and antagonism that another girl of 15 expressed when her teacher tried to make friendly advances to her S G's personality difficulties were related to his grandiose attempts to compensate for the damage that infantile paralysis had done to his physique E R, although 15 years of age, was still a baby and very self-centered E P was flat and dull and had very few social graces.

Cases included under "domestic disturbances" were children mishandled at home by parents whose own emotional lives were disorganized In the case of F M, a mother was trying to "make weight" by over-solicitude for the stepfather's repudiation of the child J. S. was able to do what he liked with his mother because of her own inadequacies E B. was an unconventional child being brought up in a conventional home environment J D. was the only boy in an Italian family of many girls H. S never had a chance in a noisy quarrelsome atmosphere after the birth of siblings R C, a boy of 9, let his vivid imagination run loose in a home which was badly broken up by a selfish stepmother and a previously alcoholic father

Back of these psychiatric diagnoses lie the more important causes of the children's difficulties Table 3 represents an attempt to organize these causes It will be seen that the number of causes does not agree with the number of cases, children having sometimes more than one assignable reason for their problems It should also be borne in mind that the diagnoses and causes of maladjustment are

TABLE 3

CAUSES OF CHILDREN'S DIFFICULTIES

Family difficulties	30 cases
Personality difficulties	17 cases
Physical factors, including heredity	16 cases
Sex conflicts	8 cases
Relationship to friends	7 cases
Changing interests	7 cases
Academic disabilities	5 cases
Bad judgment or low intelligence	4 cases
Insufficient recreational outlets	3 cases
Bad social conditions	3 cases
Miscellaneous	8 cases

not identical The diagnosis is an attempt to weigh up the causes and to pick out what appeared to be the nucleus of the problem This nucleus frequently had many underlying causes, as shown in Table 3

Undoubtedly, disturbed family interrelationships are the chief causes for children's poor school adjustment Such domestic difficulties include a subtle tie-up of feelings between various members of the family, with the consequent displacement upon the child of undesirable emotions As a result of this disturbance in the emotional life of the parents, there follow poor methods of disciplining the child at home The child's reaction to these undesirable types of discipline is carried over to the school because of the way children identify with parents teachers who stir up their resistances In this way an interplay between the child's personality and home or school takes place One-half of the children studied in this investigation suffered from markedly undesirable personality deviation. A vicious circle is made by the way the home reacts upon the child and the child reacts to these undesirable home tensions The school is, for the most part, an extension of battle ground, sometimes minimizing, sometimes intensifying what goes on at home. A study of the teachers' personalities would complete the picture

Among other physical factors listed as responsible for children's behavior problems manifested at school were included heredity and poor physical hygiene Only cases of outspokenly bad heredity were included One is perhaps a little surprised to find that the number of physical factors contributing to behavior disturbances is so large In one case sinus trouble was undoubtedly adding to the child's irritability In another case physical frailty diminished the girl's scholastic drive and her ability to concentrate Also included among physical causes of maladjustment were accidents which had led to brain damage. There were two examples of this There were also two cases where constitutional endocrinological make-up was definitely skewed. These endocrinological disturbances seemed to be related to homosexuality—present in both these cases As an example of bad hereditary influences upon behavior is the case of a girl aged 14 who was hyperactive Her younger brother had hydrocephalus, another brother was encephalitic, and a third sibling was a mental defective In another case of physical factors influencing behavior a child had had an instrumental birth with rather serious results W N was over-developed physically but poorly developed emotionally

This discrepancy, among others, heightened conflict and interfered with his social adjustment. E L. was going through the physiological changes, including sex irritation, which accompany puberty Hysterical spells and interest in older boys were the result of this endocrinological over-activity.

The children's personality traits which were contributing their share to the disturbed behavior at school included many variations. A M was a romantic individual whose thoughts turned to airplanes and adventure at times when x^2-y^2 was being discussed. J S was prosaic and superficial, money-making was his chief interest. He will probably turn out to be a successful business man, but a very poor historian E S was interested only in seeking and obtaining pleasurable returns She could stay up at night dancing until any hour, but became tired as soon as she had to study. J. D had lived a humdrum existence and felt socially inferior Her bid for recognition took the form of writing herself romantic love letters and staging artificial fainting attacks I. C had never grown up, she was generally immature. Her intellectual level was also very low. S G was happy only when in positions of power; he could not walk. As a result of the influence of older friends, J. D wanted to become a man, against his mother's wishes. E L. was self-centered. She had had her own way while her mother lived; why, she thought, should conditions be different under her stepmother's roof? E. C thought she was not good looking. Although fifteen, she did not mix with other girls or boys Instead of making friends with girls, she stole their purses

Sex conflicts were of two types. Some of the girls may have had, or desired to have, some form of sex experience and were somewhat tied up in knots about it There were some boys who had latent homosexual trends. In some cases, both among boys and girls, were found sex conflicts associated with stealing In these cases stealing may be, in the minds of children, a less serious outlet for the tensions worked up over sex conflicts

Under the heading "changing interests" are found those cases of adolescents who were trying to break away from their younger interests and engage in more adult activities. The old games and pasttimes no longer satisfied them Dancing, later hours, and older companions were now sought—much against the wishes of the parents. These cases of puberty or early adolescent dissatisfaction go hand in hand with those cases where "relationship to friends" is

held partly responsible for the client's difficulties, although in some of this latter group of cases insufficient social and recreational stimulation was also found as one of the causes of maladjustment.

One child's difficulties, namely, shyness and antagonism toward authority, were related to an IQ around 70, the child did not know how to handle her environment There were three other cases where, in spite of an adequate IQ, the child was not using good judgment. One was a college girl who could get along with no one Her poor judgment resulted from her emotional difficulties

The five cases where some academic or psychological disability rather than psychiatric disharmony existed included instances where the child's study habits were poor, where there was difficulty learning some special subject, such as mathematics, and where the grade placement was too low to stimulate the child

While all the above causes for disturbed behavior were agreed upon only after adequate study of the case, it is readily admitted that further study might uncover even more fundamental causes. In some cases only the upper layers of the difficulties were uncovered; in other cases one was able to dig more deeply There is no limit to the depth to which one can theoretically go in the matter of tracing cause and effect in behavior There is a limit, though, to the depth one need go for practical therapeutic purposes As a result of this study, it is possible to say that certain types of children's difficulties, especially many of those manifested in school by bright children, are the results of a subtle interplay between personalities those of parents and adults, e g , teachers, on the one side, and those of children, on the other Side by side with these chief influences upon behavior are certain hereditary, constitutional, and physical factors which serve to reduce the resistance of the child and make it easier for psychological excitants to operate.

Having in each case outlined what was considered the dominant, underlying bases for each child's difficulties, it was then possible to prescribe treatment directed toward changing these disturbing causes at home or at school and producing more desirable effects Where it did not seem possible to do this, an attempt was made to fill in some of the deficiencies in the child's life by a recreational program, or to stimulate the child directly through psychotherapy. The actual prescriptions, as written, can be classified as shown in Table 4.

The recreational and social recommendations were not blanket recommendations by any means. They attempted to meet the specific

TABLE 4
PRESCRIBED TREATMENT

Recreational and social changes	26 cases
Psychiatric treatment	24 cases
Academic changes	22 cases
Social service work with patients	16 cases
Instruction to adults	17 cases
Jobs	9 cases
Physical treatment	6 cases
Change of home	5 cases
Hospitalization	3 cases
Camp	3 cases
Big brother	1 case
Court supervision	1 case
Financial help	1 case

needs of each child In the case of one child who had, it was felt built up himself too much on the physical side, membership in the non-athletic clubs of his school was arranged. To the parents of this same boy, who did not like to stay at home, but who was interested in sciences, the establishment of a little experimental laboratory in his own home was suggested. In the case of a boy, aged 10, whose mother was stirring up his resistances, clubs where he could get a different and more intelligent kind of supervision away from home were prescribed In regard to J S. who was developing along such prosaic lines, membership in the school's cultural clubs, such as debating and dramatics, was suggested E B, a girl on the trail of boys, was helped to get into a mixed group. Here she would have a good outlet for her sex urge with understanding supervision.

It is interesting to observe that a recreational prescription was written in 26 out of 36 cases Does this mean that children whose recreational life is not adequate or satisfying are more likely to become difficult problems to handle; or does it mean that troublesome children are difficult to satisfy in the matter of play? Probably an interreaction operates here between cause and effect

Psychiatric help for the patient was recommended in 24 of the 36 cases In 10 of these cases it was suggested that the worker act as psychiatrist; in 7, the teacher; and in 7 others, the psychiatrist designated himself. The reason for this division of labor among psychiatrist, teacher, and worker was only in part due to the fact that the psychiatrist could not follow up all cases In some cases it was felt that encouragement should come from the school itself, as in one

instance where a boy was doing poor school work in spite of a satisfactory IQ Frequently the worker had already made a very excellent contact with adolescent girls who had no friends. The worker was instructed in such instances to try to fill in the social life of the girl through her own contact Such social occasions also gave the worker excellent opportunity for a deeper psychiatric rapport The psychiatrist took on those cases where he felt man-to-man talks would be helpful, or where, as in one or two cases, further probing would help both the study and the patient It will thus be seen that the nature of the psychiatric help, by whomever carried out, varied from encouragement to attempts to give the client some insight into his own problems.

The academic recommendations covered suggestions with regard to change of subjects, change of course, or even change of school A M was recommended for a science course instead of the academic course he was following All his drives were in the direction of engineering. E C had the reputation for being dishonest in her own school. It was felt that it would be too difficult for her to make good there, and a fresh opportunity was found for her in another school E. P was as indifferent as a girl could be toward school. She was failing in all subjects, and, through her school contacts, was coming closer and closer to sex delinquency It was recommended that she leave school and take up a business course. This she did with happy success. E L., who was failing in school because of a post-traumatic psychosis, was allowed to remain in her present school at her own pace W. N., a boy with an IQ of 138, needed a special school where he would have a chance to progress as rapidly as he is capable of doing and where he would have a varied, selective curriculum to stimulate him. E B, a girl of 13, is only in the fourth grade in spite of an IQ of 103 She has been penalized because of her foreign parentage A little special help in language will fit her for more rapid advancement E G. is much brighter than the teachers think she is Her IQ is 112 She was advanced a grade on the suggestion of the clinic, and both her work and conduct improved considerably.

In 16 cases it was recommended that work be done with the parents by the social workers. This work was independent of the instruction given directly to parents by the psychiatrist where it was felt that more intensive treatment was not needed. The work done with the parents included attempts to change their methods of

handling their children, or attempts to give them insight into *their own problems*, and thus improve their relationship to their children.

In 9 cases it was recommended that children should replace school by work or supplement school by work, in some cases during the summer and in some cases after school. While a "change of home" was recommended in 5 cases, such a change would not necessarily mean a permanent foster home. A change of home might be temporary, as during the summer months, or for a trial period. It might mean only a prolonged visit to relatives.

In one case where a boy seemed to respond only to a big stick, it was recommended that court supervision be continued. In another case, it was thought wise to lighten the burdens of a college girl by a little financial help from the clinic. The physical recommendations included specific medical treatment or change in physical hygiene.

For various reasons these recommendations were not always carried out. Table 5 estimates how frequently the prescriptions were dispensed by the social druggists.

Social and recreational suggestions were carried out only 50% of the time. The reasons for this are not hard to find. Frequently the school had not the type of club or recreational service suggested. The community, too, did not always supply what was demanded. In some cases the patient or the patient's parents balked at help.

It is rather encouraging to learn that over 70% of the recommenda-

TABLE 5

ESTIMATES OF HOW FREQUENTLY TREATMENT WAS CARRIED OUT

	Treatment prescribed	Treatment carried out
Recreational and social changes	26 cases	13 cases
Psychiatric treatment	24 cases	23 cases
Academic changes	22 cases	16 cases
Social service work with patients	16 cases	15 cases
Instruction to adults	17 cases	8 cases
Jobs	9 cases	3 cases
Physical treatment	6 cases	4 cases
Change of home	5 cases	1 case
Hospitalization	3 cases	3 cases
Camp	3 cases	2 cases
Big brother	1 case	1 case
Court supervision	1 case	1 case
Financial help	1 case	1 case

tions with regard to change in school procedure were carried out This points to very excellent cooperation on the part of the schools. It is not always easy for schools to permit a boy to change his program solely on the advice of an outsider In this connection, it is probably only fair to point out that the clinic was unusually assured of the school's cooperation before study was undertaken, and that the clinic's workers were usually tactful and flexible

Direct work was done with the patients' parents in almost all cases where it was recommended Such direct treatment was carried out by the clinic's own social workers who, of course, would be expected to carry out the clinic's recommendations. The intensity of this treatment of parents does not show up in this report One feels that, because of the rather heavy case load carried by workers, treatment was not always as intensive as was recommended This same criticism applies to the psychiatric treatment, especially that done by the psychiatrist It was not always possible to see the subjects as frequently as one would like. The children were usually a long way from the clinic and could not always leave school with perfect facility This mechanical difficulty raises the question as to the best location for a clinic working with schools Should the clinic be established right in a central school building?

It is interesting to note that only in one case out of the five in which a change of home was recommended was this recommendation carried out. This success occurred in the instance of an adopted boy living with relatives who did not want him The boy, too, was anxious to get back to his brother with whom he felt more secure The fact that a recommendation of foster-home placement was not carried out in the four other cases demonstrates to some extent the resistance of the parents to any change in the existing emotional status of the case.

The next step in the investigation was the attempt to evaluate the results of treatment. This is a very important procedure

A clinic's own personnel (for the purposes of self-education) and those interested in mental hygiene in one way or another are interested in learning just how successful are treatment policies The following summary is an attempt to estimate the results of the year's work from the point of view of benefit to clients. In some cases, of course, it is too soon to learn what improvement may later manifest itself, in other cases it is likewise too soon to learn what regression may occur. The bases used for judging improvement or

TABLE 6
EVALUATION OF RESULTS OF TREATMENT

	Number of cases
No. improvement 0% improvement	2
Little improvement 20% improvement	4
Much improvement 40% improvement	13
Very much improvement 60% improvement	6
Complete recovery 100% improvement	2

lack of improvement were the reports from all possible sources—teacher, clinic worker, parent and the psychiatrist's re-evaluation in cases where he re-examined the patient

For the purpose of a rough-and-ready statistical estimate of the amount of improvement shown by any one case a 5-point scale was adopted The cases are measured on this scale according to whether they showed no improvement, little improvement, much improvement, a very great deal of improvement, and complete cure Side by side with each descriptive class a mathematical unit of percentage of improvement was added, varying from 0% to 100%. Table 6 shows how cases were re-evaluated after treatment in terms of the percentage of improvement subsequently displayed

It will be seen from this table that two cases failed to improve at all, these were balanced by two cases which showed complete recovery. The cases which failed to improve at all were one case of dementia praecox, and one case of an encephalitic condition. In both these cases the nature of the problem was too serious and fundamental to be helped in a short time. One of the cases showing 100% improvement, as estimated by the school, was a boy whose mother claimed he was neurotic, she was afraid that he was becoming homosexual. By means of a big brother—who, by the way, was really a big sister, namely, the school counselor—this boy was introduced into social life under encouraging circumstances. He blossomed out socially, and his school work improved at the same time. The other

case of a so-called complete cure is worthy of mention This girl of 18 had few friends and had been diagnosed on two other occasions as a dementia praecox case for whom, as usual, a bad prognosis had been given She was seen again at the beginning of the year and her condition at that time seemed to resemble a manic-depressive psychosis. At the present time, due to the stimulating effect of the clinic and the cooperation of the school, she is able to take a job. Her parents report that she is much more social at home, too. Moreover, the psychiatrist feels that she is not likely to slide back into her old condition.

There are many other cases where encouraging results were obtained There was one interesting case of a girl suffering from fainting attacks. After she had been studied by the clinic the spells disappeared and the school reports "that J. D has done good work and made progress" A. B was referred because of forgery and truancy His teachers now report "that he is a completely changed boy and there has been no dishonesty in a couple of months." D L, who engaged in fantastic lying and made bids for attention, showed improvement for a number of months and then had a slump just before her examinations, to escape from which she invented an attack of laryngitis

One frequently found that the initial symptoms for which the child was referred to the clinic disappeared, although more serious underlying character defects remained This change occurred in one case of a girl who had crying spells She gave up her crying, but was still very difficult to manage. Another girl who was referred for stealing did not steal again after the clinic saw her, but subsequently lied herself out of a job.

S. K., whose problems were serious enough to require removal from home and hospitalization, is now "quiet and pretty well adjusted" "There is progress in the home situation" where previously marital disorder was producing the nervous wear and tear to bring about this boy's problems H S is a "perfect gentleman" at a camp away from home, but his gentlemanliness disappeared in part after a few days at home. E. L., who was referred for stealing and antagonistic behavior toward the teachers, has stopped stealing, but "there is no improvement in school." M. W., who is a truant and uncontrollable, "is now making better effort Truancies less frequent Appearance better." R. C was referred because of stealing and overactivity. His stealing has disappeared and he is much quieter. He has been kept in bed because of illness Perhaps he has not had much op-

portunity to steal there! E H, who was "disorderly in class and cruel to younger children," showed "a good deal of improvement." Scholarship improved and his behavior was far less annoying in school. Unfortunately, he became more cruel at home and attempted to choke his mother. She demanded his removal from the home S. K, who was "restless, violent and unstable at school" and also a truant, has manifested less truancy, improved in his school work and "his appearance is 100% better"

While these results seem fairly encouraging, one has to be cautious about putting a final stamp of approval upon them. It is quite likely that should the clinic withdraw its supervision or interest in these cases, some of them would revert to their previous unsatisfactory status On the other hand, it is quite possible that a few of the cases where the results are so far unsatisfactory will show some improvement as further intensive work is done with them

It is interesting to speculate about what caused the improved behavior in our clientele While, of course, successful methods of treatment with one case would not work with another, it is nevertheless worth while learning what, by and large, constitute the main factors making for successful therapy The analysis shown in Table 7 is an attempt to estimate these factors

Under "clinic's efforts" were listed cases which could not presumably have improved without the clinic's intervention This intervention took various forms It might consist of work with the parents For example, R. C was kicking up because of a feeling of insecurity which came from the tangled emotional mess of a father-stepmother interrelationship The clinic's work with the parents had a distinctly favorable reaction upon the child The clinic's efforts might be part of a process of encouragement and stimulation for the

TABLE 7

ANALYSIS OF FACTORS MAKING FOR SUCCESSFUL THERAPY

Clinic's efforts	15 cases
Change of parental attitudes	12 cases
School cooperation	9 cases
Recreational resources	5 cases
Academic changes	4 cases
Change of home	3 cases
Insight acquired by clients	3 cases
Work	2 cases
Influence of court	1 case

patient, as in the case of the extremely depressed girl, where encouragement by and opportunity for social contacts with clinic workers undoubtedly helped to bring the patient out of the hole Sometimes the clinic's efforts consisted in interpreting the patient to the parents, as happened with E B, whose parents were trying to mold her into a rigid routine Occasionally the clinic, through the psychiatrist, provided the successful touch This happened with A B, a latent homosexual, who was relieved of his conflicts partly by the psychiatrist. W. N, who felt superior to life, was helped to obtain a new standard of values by means of insight the psychiatrist helped to give him There is the interesting case of J D, who was heading for the stage but riding rough shod over the world in the meantime Because of his supercilious attitude, he turned out to be one of the most difficult personalities with whom the present examiner ever had to deal The examiner tried "shock therapy" upon this boy. He told the boy as frankly as he knew how a few disagreeable truths about himself The boy flared up, stampeded out of the office, bawled out the teacher for sending him to the psychiatrist—and improved in his work and attitude! Undoubtedly, a change of attitude on the part of parents helped along this adjustment; but through the psychiatrist's attitude the boy was made to face himself very clearly for the first time He evidently did not like what he saw in the mirror that was held up to him, and set about to change the picture E. B, a psychopathic personality, was undoubtedly helped by the clinic *in many ways* Financial help from the clinic kept her in college, and the worker's close sympathetic supervision gave this girl practically her only friend.

Twelve cases were improved because of a "change of attitude on the part of the parents" This type of therapy, fundamental in child guidance work, is one of the most difficult to administer, inasmuch as the causes for the original parental attitudes are so deep-seated It is easy to understand how parents who have been handling their children in a certain way for 15 years are not, as a rule, likely to change their tactics very readily as a result of outside intervention; such parental methods are usually too well built upon a very close emotional need It is rather outstanding, that 12 parents did mend their ways to some extent—even if some of them regressed later The initial attempt to treat them has been successful, and chances for even larger success subsequently are greater because of these initial fortunate contacts.

In the 9 cases which were helped by the school this assistance came

largely through the clinic's insight into the children's difficulty. This insight enabled the worker to act as liaison officer between the child and the teachers, and interpret the child's needs to the latter successfully Without the cooperation of the school, however, and a change of attitude by its personnel, a happy prosecution of the case would not have been possible

Four children were distinctly benefitted by improved school rating

Five children were found fresh recreational outlets; the stimulation they received from games carried over to their school work. One seldom, of course, found these recreational satisfactions acting alone as the basis for the children's improvement in general; usually they went hand in hand with improved methods of handling by parents, and prestige at home for the child

Three children readjusted themselves away from home, their problems practically disappearing in their new environment.

Two adolescents got along well at work, at school they had caused much trouble

When the chief causes of bringing about a better adjustment of children, namely, clinic's efforts, work with parents, and school's cooperation, are analyzed a little more carefully, it will appear that the basis of this improvement in children's behavior is the effect of other personalities upon the children, whether it be that of the teacher, psychiatrist, clinic worker, or, more important, parent This finding fits in with that arrived at as to chief causes for children's difficulties, namely, the influence of the home, where again the personality of parents is the primary toxin.

Equally important as a study of causes for success in the matter of therapy is a study of causes for lack of success. An effort was therefore made to estimate the reasons why the clinic failed partially or

TABLE 8

REASONS FOR FAILURE OR PARTIAL FAILURE

Inability to reach parents	4 cases
Severity of the disorder	3 cases
Poor contact with patient	2 cases
"Too many cooks"	2 cases
Personality of personnel	1 case
Low-grade family	1 case
Inadequate study	1 case
Treatment not intensive enough	1 case
Unemployment	1 case

completely to readjust its clients The results of this study are outlined in the list of reasons for unsuccessful therapeutic efforts shown in Table 8.

In the same way that personality is a cause of success in treatment, it would appear that personality is equally a cause of failure Out of the above 14 cases where success was poor, 9 of them appeared to have suffered from a personality in contact with the case, either in the home, or the clinic, or the school The personality difficulties of those of us working with these cases which interfere with therapy are much the same type of problem as unsuccessful parents show Our own difficulties consist of resistances and blockings and emotional identifications which prevent us from retaining our objectivity and doing what is best for the patient, irrespective of our own feelings. From these findings as to the cause of unsuccessful therapy, it would appear that when very severe, practically hopeless cases are eliminated from the clinic's intake by a more careful admitting arrangement, a well-adjusted staff working with patients who are at least capable of readjustment should produce most excellent therapeutic results.

The case requiring more intensive treatment before better results could be obtained was that of a highly emotional girl requiring a frank psychoanalysis to break up her emotional immaturities She improved while she thought she had the attention of the psychiatrist, but broke down when the stress of reality became too severe at school

Failure in one case was due to the fact that not enough was known about the situation before treatment was started. The girl herself was contributing a good deal to her problems whereas the clinic tended at the outset to place too much emphasis upon the home

One case of a very low-grade family was just impossible at the outset from every point of view and little could be expected from it.

In two of the three cases where "severity of the disorder" was responsible for failure, permanent brain damage had taken place. No amount of child guidance work can replace damaged brain cells. The third instance in this group was suffering from dementia praecox.

One case was not improving because of the clinic's inability to find a job for a discouraged, low-grade girl

"Too many cooks." It is worthy of mention that a clinic outside the school system collaborating on mental hygiene work with so many school officers, many of them in positions of authority, felt that in only one or two cases did this liaison arrangement definitely interfere with the client's welfare It is quite possible, though, that if the

clinic had been able to control the total situation, many other cases would have manifested even greater improvement. There is little doubt that, by and large, study of a case and therapy—it is frequently impossible to draw the delimiting line between these two—should be concentrated in the fewest possible hands.

As another independent effort to evaluate a mental hygiene approach in the study and treatment of children's scholastic problems, the schools which had referred the above-mentioned cases were surveyed by means of a questionnaire and a personal interview by the clinic's workers. Representatives of these schools—principals, vocational counselors, student advisers, psychologists—were first asked: what help did they expect to receive from a child guidance clinic. Answers to this question showed marked agreement. It was the function of the clinic to make "a thorough investigation of the child's problems." Schools expected "definite constructive suggestions as to treatment of the problem." From the clinic's study, schools wanted to obtain insight into and understanding of the child's behavior. One or two schools thought that a good physical examination was an essential part of the clinic's activities. (A physical examination was done on almost all the above-mentioned cases.)

The schools asked what they felt a child guidance clinic had actually given them replied that they thought they could accept the results of examinations with confidence. These satisfactory examinations—psychological, physical, psychiatric, together with social history—had given them insight into the child's difficulties. On the basis of this insight and the specific therapeutic suggestions, they had been helped in adjusting the difficulties for which the children had been referred. One prominent school principal expressed her reactions as follows: "When we get to a blank wall, a child guidance clinic is where we turn to. We don't get so discouraged any more, knowing there is help and understanding to be had."

All the schools agreed with more or less emphasis that the behavior of the children had improved as a result of the clinic's cooperation. It was interesting to learn that this improvement had been brought about from the school's point of view in four ways:

a. The school had received broader understanding of the child's difficulties.

b. The examinations and the worker's contact with the school in interpreting these examinations had brought the teachers together

and had helped them to cooperate in over-coming the child's difficulties

c Many schools thought that the physical care which the clinic had given through its workers was helpful in adjusting the child

d. School principals seemed to feel that through the worker the parents' cooperation in working out the clinic's therapeutic suggestions had been obtained.

Almost all schools agreed that they had found no difficulty in carrying out the clinic's suggestions Only occasionally did the school machinery interfere with the clinic's therapeutic program. Evidently, the principals had the power, as one teacher expressed it, "to cut red tape in the school and see that plans are carried out" Most schools "welcomed any suggestions and were glad to carry them out. The faculty as a whole was cooperative and helpful."

Asked how the clinic might have been more valuable in handling the schools' cases, principals answered in chorus "Give us more of you" The schools seemed especially anxious to have assigned to them "a capable worker to give more time to this school's work"

The clinic was anxious to learn, too, whether it had been successful in educating the teachers to become psychiatrically minded so that there "would be a favorable carry-over to other cases not studied by the clinic." From the answers obtained to this question it seemed that, by and large, the schools obtained a much better insight into a mental hygiene approach Some schools "used the clinic's reports as guides to solve other problems" "The teachers' enthusiasm has been stirred up and by and large they are anxious to learn more about child guidance investigations."

The clinic was also interested in the following administrative problems.

Did the schools think that the clinic's work would be more effective if the clinic visited the school, instead of having the children come to the clinic's offices? Opinion was divided on this point. Feeling was a little stronger, however, in favor of the view that a child guidance clinic would do better work if operating in the school building Schools felt that this arrangement would overcome parents' prejudice against social agencies For the most part, the principals felt that it was easier for schools rather than clinics to get the cooperation of the parents Parents were flattered by the school's attention. Unfortunately, most of the schools felt they had not the space in which to house

a clinic One school psychologist, however, believed that in going to an outside clinic, the child would be more likely to feel that he himself was the one who was going to profit from the study, whereas an intramural study would emphasize in the child's mind the benefits likely to accrue to the school. All schools agreed that in studying the home conditions behind children's problems the emphasis should be laid upon the training and tact of the person carrying out the home investigation, rather than upon the affiliations of the worker The schools agreed, of course, that it was necessary to use social agencies at the present time because no member of their own personnel had time or training for careful detailed social service work.

Most of the schools were interested in having lecture courses for their staff, especially where such courses would bring academic credit to the teachers Some of the principals seemed to feel that the teachers were rather swamped with lectures now and might not have sufficient time to attend mental hygiene courses Those teachers who were already partly trained in a mental hygiene approach felt they could get the most benefit from small group conferences under competent leadership rather than by attending popular and overcrowded lecture courses

From a year's intensive study of 36 cases of scholastic maladaptations referred to a child guidance clinic, and from reports received from the schools making use of the clinic, the following conclusions can be drawn

The chief reason discovered for the schools' desire for a clinical investigation of a problem case was the fact that the pupil was interfering in one way or another with school routine. Clinical study showed that personality disturbances among the children, family difficulties, and physical influences were easily the major causes of the children's apparent scholastic difficulties. In this connection it is interesting to observe how home influences can play a large part in producing scholastic problems.

In attempting to overcome the child's maladjustment a large number of recommendations were made with regard to change in academic programs. Many prescriptions were handed out calling for recreational and social changes Psychiatric and social service work with patient, parents, and teachers took care of even a greater number of prescriptions By and large, the schools were very cooperative in putting through the academic changes recommended

In the matter of results obtained from the clinic's work it was

possible to estimate that cases studied and treated had shown 50% improvement in behavior. It was also estimated that the causes for successful treatment were largely the clinic's efforts in producing a change of parental attitudes, and work with the schools in carrying out the recommendations of the examiners The chief causes for lack of success were the severity of the disorder, or inability to reach parent or patient

The schools, when questioned, seemed to feel that they had received a good deal of help from a child guidance clinic They felt satisfied that the clinic's investigations were thorough, and gave teachers insight not only into their own pupils' difficulties, but insight into the motivation of disturbed behavior more generally Principals expressed themselves very emphatically as to the need for an extension in the schools of child guidance work. School authorities also stated that they are interested in discussions and conferences on children's behavior problems in general

Columbia University
New York City

UNE ÉTUDE CLINIQUE DE L'APPLICATION DE L'HYGIÈNE MENTALE AU TRAITEMENT DES PROBLÈMES SCOLAIRES DES ENFANTS

(Résumé)

Les écoles publiques s'intéressent surtout à faire suivre une routine aux enfants. Quand un enfant résiste aux meilleurs efforts du personnel de l'école, on demande à une clinique psychiatrique pour les enfants d'aider dans ce processus d'adaptation. Apres avoir revu l'étude pendant une année et le traitement des enfants à désadaptations scolaires, une clinique psychiatrique a appris que les difficultés de personnalité et les incapacités physiques chez les enfants, et les désadaptations émotives chez les parents ont été les principales causes de l'insuccès à l'école

Les principales méthodes employées pour surmonter les difficultés qui empêchent le développement des enfants à l'école ont compris des programmes de récréation, des changements de milieu et du travail psychiatrique avec les enfants, les parents et les maitresses d'école On a trouvé que les meilleurs résultats ont été obtenus par l'influence de quelque personnalité, celle du psychiatriste ou celle du travailleur social sur l'enfant, le parent ou la maitresse. La personnalité semble avoir une plus grande influence formative sur les enfants que l'organisation scolaire ou n'importe quel autre agent impersonnel, par ex, la récréation

En somme on a pensé que la clinique a réussi de cinquante à soixante pour cent dans ses efforts de faire adapter les enfants traités à l'école

Les principales causes du manque de succes ont été le gravité du problème de l'enfant et la difficulté d'influencer les parents par son approche psychiatrique

LEVY

EINE KLINISCHE UNTERSUCHUNG DER ANWENDUNG DER GEISTIGEN HYGIENE AUF DIE SCHULSCHWIERIGKEITEN VON KINDERN

(Referat)

Die offentlichen Schulen interessieren sich besonders fur die Einfugung der Kinder in eine Routine Widerstrebt ein Kind den besten Bestrebungen des Schulpersonals, so wird eine Kinderfuhrungsklinik (child guidance clinic) zur Hilfe bei diesem Bildungsvorgang herangezogen Eine gewisse Klinik dieser Art hat aus nochmaliger Durchsichtigung einer jahrlangen Untersuchung und Behandlung von Kindern, die sich der Schule schlecht anpassten, erfahren dass Karakterschwierigkeiten (personality difficulties) und physische Unfahigkeiten (disabilities) bei den Kindern und affektive Storungen (emotional maladjustments) bei den Eltern die Hauptursachen ungenugenden Schulfortschrittes darstellten.

Die Hauptverfahren, die zur Überwindung der die Ausbildung der Kinder storenden Hemmungen (handicaps) verwendet wurden, bezogen sich auf Erholungsprogramme (recreational programs) Anderung der Umgebung, and psychiatrische Arbeit mit Kindern, Eltern, und Lehrern Man fand, dass die besten Resultate durch die Einwirkung des Karakters (personality) des Psychiaters oder des Sozialarbeiters (social worker) auf Kind, Eltern, oder Lehrer erzielt wurden Der Karakter scheint auf die Kinder einen starkeren Bildungseinfluss auszuuben, als die Schuleinrichtung oder irgend eine andere unpersonliche Einwirkung wie zum Beispiel die der Erholung

Im grossen Ganzen fuhlte man, dass die Klinik, in ihren Bestrebungen, eine gute Schulanpassung (successful school adjustment) bei den von ihr behandelten Kinder zu erzielen, einen Grad von 50 bis 60% in Bezug auf Erfolg erreichte Die Hauptursachen des Mislingens der Klinikarbeit lagen in der Bedenklichkeit der durch das Kind gestellten Aufgabe und in der Schwierigkeit der Erreichung der Eltern mit der psychiatrischen Annäherungsweise.

LEVY

SHORT ARTICLES AND NOTES

EDITED BY
John E Anderson
Charlotte Bühler
Cyril Burt
Toshio Nogami
Ivan P. Pavlov
H. Piéron
Calvin P Stone

SOME NEW BASES FOR INTERPRETATION OF THE IQ—ERRATUM

Beth Wellman

Figure 1 of the article entitled "Some New Bases for Interpretation of the IQ" as it appears on page 117 (No 1) of this volume is up-side-down The accompanying figure shows the figure in its correct position

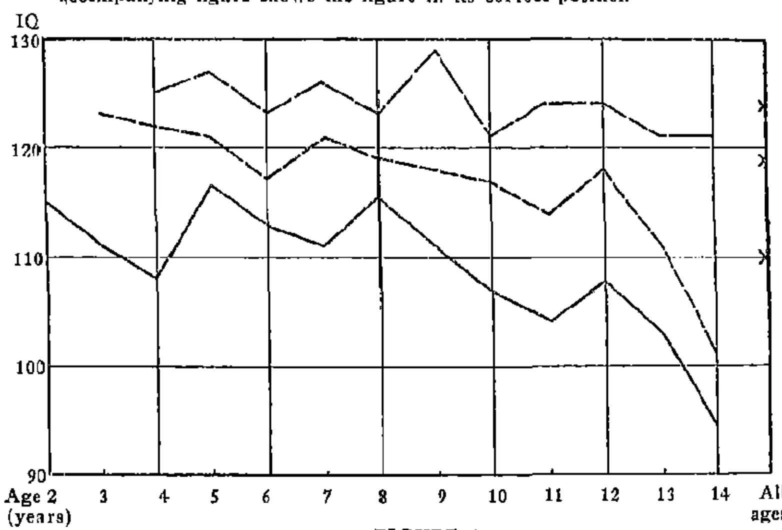


FIGURE 1

Mean IQ on Repeated Tests

1st test ————, 2nd-3rd test — — — —; 4th-7th test —— —- ——

Number of Children

Age (years)	2	3	4	5	6	7	8	9	10	11	12	13	14	All
1st test	93	172	150	176	180	95	72	69	77	67	71	50	61	1333
2nd 3rd test		112	181	162	137	107	71	53	54	60	40	30	20	1027
4th 7th test			39	98	79	35	43	52	53	42	55	46	32	574

THE NON-RANDOM CHARACTER OF INITIAL MAZE BEHAVIOR—ERRATUM

WAYNE DENNIS AND R H HENNEMAN

Lines 10 and 11 on page 400 of this journal, Volume 40, No 2, June, 1932, were printed in a transposed position, their proper position being between lines 16 and 17 of that page The section should read as follows

The record of this first group as a whole was only slightly less remarkable than that of the two rats just discussed Two rats made only one error each, two more made only two each, a total of 11 rats each made no more than 4 entrances into culs-de-sac The greatest number of total entrances into blind alleys was 10, while the average number was 4 5 In other words, not a single rat entered all culs-de-sac on the first run, and the average rat entered only one-third of them

However, we are chiefly interested not in the total errors but in the direction taken by the rat in leaving each unit of the maze *for the first time*

LEARNING OF A PETZ'S CONURE

R H MASURE

Much has been said about the intelligence of parrots, many authors have even suggested that because of their superior intelligence the parrots, family Psittacidae, should be placed at the top of the class Aves. Although some work has been done on avian learning, practically nothing has been done that shows anything definite about the intelligence of any sort of parrot. I am presenting this work here, which was done under Dr. W C Allee, as a small contribution to this unexplored field and as a stimulus for further exploration

The experimental work was done on a Petz's conure (*Eupsittacula canicularis*, Linn), a representative of the genus Eupsittacula, which is one of the largest genera of the sub-family Psittacinae, the typical parrots The conures are native to Central America, from Mexico to Costa Rica, where they abound in large flocks This particular bird was exceedingly tame I had kept it at home for several months previous to these tests, and during that time he became so tame that he followed me around much like a dog will follow its master I am inclined to believe that the bird was a male, although I cannot be sure as it is quite difficult to determine sex of parrots by external characteristics

PROCEDURE

The first experiment was conducted in a two-choice box (Figure 1), with light and darkness as the two stimuli The bird was to choose the dark-

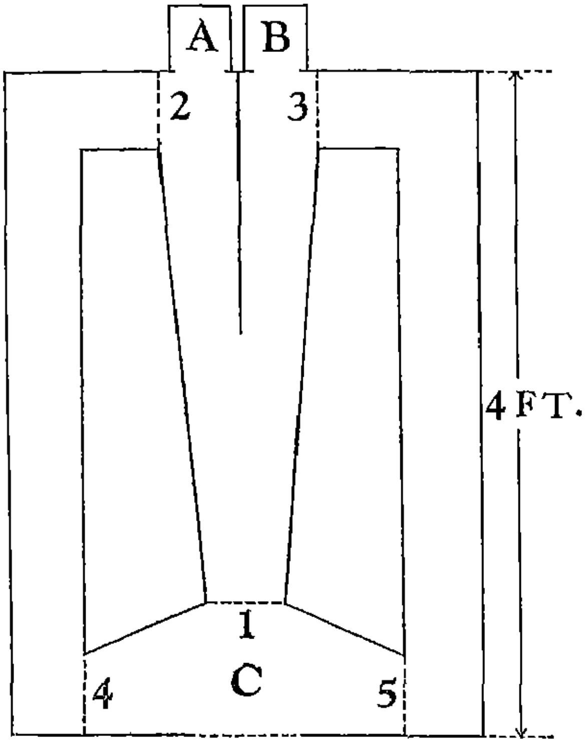


FIGURE 1

Two-Choice Box

1, 2, 3, 4, and *5*—movable doors manipulated from the front of the box
A and B—boxes containing 15-watt Mazda lamps shielded by ground-glass circles 2½ inches in diameter.
C—the vestibule

The box was shielded at the front so that the investigator could not be seen by the bird. The bird was placed in the vestibule and door *1* was opened allowing the bird to pass into the area of choice The door at the end of the correct alley was opened and the door at the end of the incorrect alley was closed. After the bird passed door *2* or *3*, as the case might be (the stimuli were alternated from left to right), the corresponding door, *4* or *5*, was opened and the bird allowed to pass into the vestibule. Each door was closed as the bird passed it

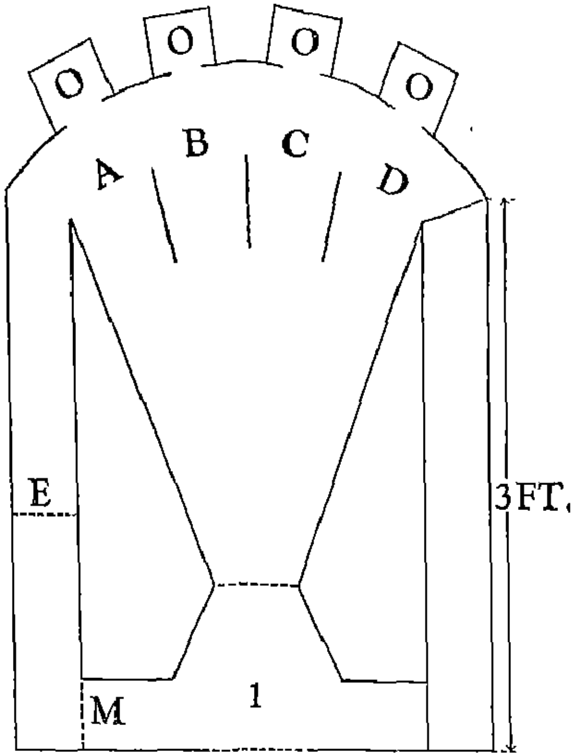


FIGURE 2

FOUR-CHOICE BOX

A, *B*, *C*, and *D*—alleys equipped with movable doors and manipulated from a large panel in front of the box that shielded the observer from the bird

M and *E*—movable doors

O—boxes containing 15-watt Mazda lamps shielded by ground glass

I—the vestibule.

The bird was placed in the vestibule and allowed to pass into the area of choice, if the incorrect choice was made, door *E* was closed and the bird was obliged to retrace his course and make another choice When the correct choice was made doors *E* and *M* were opened allowing the bird to pass into the vestibule, after which *M* was closed

ened alley, as 10 preliminary trials showed that the bird was photopositive. Twenty trials were given daily, 10 in the afternoon and 10 in the morning. Two successive days without error was considered as the criterion for mastery of the problem. As a reward for choosing the correct alley and returning to the vestibule, the bird was released from the box, allowed to crawl upon the experimenter's hand, and given a few bites of banana. The bird was adversely stimulated by the dim light and close walls of the box and made efforts to escape from this situation, therefore, release from the box was a reward, and the banana, which the bird was willing to eat at any time, was an added reward.

The second experiment, conducted in a four-choice box (Figure 2), was run to condition the bird to choose the dark alley, which was constantly shifted. At first the bird was allowed to return to the vestibule (all doors being kept open) even if he did not make the correct choice. If he made the correct choice he was released and given some banana, but if he made an incorrect choice he was given no banana and was confined in the vestibule for several moments and then made to run again. This did not prove satisfactory as the bird continually chose door *A* even though he was not rewarded. To prevent this a door was inserted at *E*, which was closed if the wrong choice was made, compelling the bird to retrace his steps and make the correct choice before he was allowed to enter the vestibule and be rewarded. This also proved faulty, as the bird had acquired a position habit and continually chose door *A* first and then made the correct choice. Door *A* was then shut, making a three-alley choice box. When this was mastered door *A* was opened. Twenty trials were given daily, and two successive days of trials without error was the criterion for mastery, as in the first experiment.

The third experiment was in the two-choice box (Figure 1), using the same method as in the first experiment, but substituting red and green light for light and darkness as the two stimuli. Ordinary colored glass plates, which were known to transmit mixed wave lengths of light and to differ in intensity of transmitted light, were placed in front of the white lights. As the preliminary tests showed that the bird was more positive to the red (8 out of 10 trials), the green light was used as the correct choice.

Tests for retention of the two-choice light-and-dark problems were made 10 and 20 days after the problem had been mastered. During this time the bird was learning the four-choice light-and-dark problem. Another retention test was made 27 days after the problem had been mastered, during the latter 7 days of which the bird was learning the two-choice red-and-green problem.

When the bird was not in a choice box he was kept in a cage and allowed as much sunflower and hemp seed as he wished to eat.

Results

TABLE 1
Results of the Two-Choice Light-and-Dark Problem

Day	Trials	Time (min-sec)	Errors
1 Morning	1	4-46	2up 3L
	2	3-24	3up 6L
	3	2-20	4L
	4	0-37	1dr 1 2L
	5	0-10	
	6	0 50	2L
	7	0-52	1dr 2
	8	0-18	1up 1L
	9	0-39	
	10	0-59	1L 1up
Afternoon	1	0 21	1L
	2	1-19	4L 1up
	3	0-12	
	4	4-20	2L 1up 4dr 2
	5	0-26	
	6	2-31	1L 1up 4dr 2
	7	1-7	3dr 1
	8	1-6	1L 2dr 2
	9	0-45	1dr 1
	10	0 29	1L
2 Morning	1	1-20	2L 1dr 2
	2	0-25	1L 1up
	3	0 34	1L 1up
	4	0-23	
	5	0-29	1L
	6	1-14	2dr 3
	7	0-18	1L
	8	0-12	
	9	0-21	1dr 3
	10	0-10	
Afternoon	1	0-13	
	2	0-9	
	3	0-13	
	4	0-12	
	5	0-20	1L
	6	0-8	
	7	0-8	
	8	0-20	1L
	9	0-7	
	10	0-9	
3 Morning	1	0-8	
	2	1-20	1L 2dr 3
	3	0-9	

TABLE 1 (*continued*)

Day	Trials	Time (min.-sec.)	Errors
	4	0-7	
	5	0-8	
	6	0-14	1dr. 3
	7	0-8	
	8	0-16	1L
	9	0-8	
	10	0-7	
Afternoon	1	0-13	
	2	0-10	
	3	0-9	
	4	0-8	
	5	0-7	
	6	0-7	
	7	0-9	
	8	0-8	
	9	0-8	
	10	0-7	
4 Morning	1	0-7	
	2	0-9	
	3	0-7	
	4	0-8	
	5	0-7	
	6	0-8	
	7	0-6	
	8	0-8	
	9	0-6	
	10	0-8	
Afternoon	1	0-8	
	2	0-11	
	3	0-6	
	4	0-8	
	5	0-7	
	6	0-7	
	7	0-6	
	8	0-8	
	9	0-7	
	10	0-7	
5 Morning	1	0-6	
	2	0-8	
	3	0-7	
	4	0-8	
	5	0-7	
	6	0-7	
	7	0-6	
	8	0-6	
	9	0-6	
	10	0-6	

TABLE 1 (*continued*)

Day	Trials	Time (min.-sec.)	Errors
Afternoon	1	0-8	
	2	0-7	
	3	0-6	
	4	0-6	
	5	0-6	
	6	0-7	
	7	0-6	
	8	0-6	
	9	0-7	
	10	0-6	

Designations of errors are as follows:
L—choice of the lighted alley
Dr.—attempt to retrace path to closed door, indicated by the number following "dr."
Up—attempt to climb the partition between the two alleys
The number in front of the letter indicates the number of times the error was repeated

These results show 100 trials for mastery of the problem, with an average time of 26.54 seconds per trial and an average error of .86 per trial. The running-time was reduced from the first day's average of 1 minute and 22 seconds to the last day's average of 6.6 seconds. The average daily error was similarly reduced from 2.8 per trial to 0.

All of the retention tests showed no error, the first (10 days) an average time per trial of 8.4 seconds, the second (20 days), 10.4 seconds, and the third (27 days), 8.4 seconds.

TABLE 2
RESULTS OF THE FOUR-CHOICE LIGHT-AND-DARK PROBLEM

Day	Trials	Time (min.-sec.)	Doors chosen	Dark door
		Door A open		
1 Morning	1	1-23	A-C-A9-B	B
	2	3-56	A4-D	D
	3	0-8	A	A
	4	2-54	A2-D2-C	C
	5	1-8	A2-B	B
	6	0-29	A2-B	B
	7	0-27	A-D	D
	8	0-7	A	A
	9	0-55	A-D-C	C
	10	1-21	A-D2-B	B

TABLE 2 (*continued*)

Day	Trials	Time (min-sec)	Doors chosen	Dark door
Afternoon	1	0-37	A-C3-B	B
	2	0-37	A-B2-D	D
	3	0-7	A	A
	4	0-44	A-D-C	C
	5	1-20	A-D2-B	B
	6	0-49	A-B	B
	7	0-46	A-D	D
	8	0-10	A	A
	9	0-32	A-C	C
	10	0-35	A-B	B
		Door A closed		
2 Morning	1	0-19	to A-D	D
	2	0-30	to A-B	B
	3	0-23	to A-C	C
	4	0-9	B	B
	5	0-29	B-D	D
	6	0-25	C	C
	7	0-9	B	B
	8	0-48	B-D	D
	9	0-28	C	C
	10	0-17	D	D
Afternoon	1	0-14	D	D
	2	0-10	B	B
	3	0-21	C	C
	4	0-14	B	B
	5	0-15	D	D
	6	0-11	C	C
	7	0-10	B	B
	8	0-41	B-D	D
	9	1-3	D2-C	C
	10	0-14	B	B
3 Morning	1	0-17	D	D
	2	0-12	B	B
	3	0-23	C	C
	4	0-15	B	B
	5	0-18	D	D
	6	0-11	C	C
	7	0-17	B	B
	8	0-19	D	D
	9	1-10	B-C	C
	10	0-9	B	B
Afternoon	1	0-24	B-D	D
	2	0-10	B	B
	3	0-13	C	C
	4	0-10	B	B
	5	0-12	D	D

TABLE 2 (*continued*)

Day	Trials	Time (min-sec)	Doors chosen	Dark door
	6	0-10	C	C
	7	0-12	B	B
	8	0-12	D	D
	9	0-14	C	C
	10	0-10	B	B
4 Morning	1	0-14	D	D
	2	0-17	B	B
	3	0-16	C	C
	4	0-13	B	B
	5	0-16	D	D
	6	0-10	C	C
	7	0-11	B	B
	8	0-12	D	D
	9	0-9	C	C
	10	0-18	B	B
Afternoon	1	0-14	D	D
	2	0-10	B	B
	3	0-10	C	C
	4	0-16	B	B
	5	0-15	D	D
	6	0-10	C	C
	7	0-8	B	B
	8	0-10	D	D
	9	0-12	C	C
	10	0-11	B	B
5 Morning	1	0-15	D	D
	2	0-15	B	B
	3	0-11	C	C
	4	0-15	B	B
	5	0-15	D	D
	6	0-11	C	C
	7	0-10	B	B
	8	0-11	D	D
	9	0-13	C	C
	10	0-14	B	B
Afternoon	1	0-17	D	D
	2	0-11	B	B
	3	0-33	B-C	C
	4	0-12	B	B
	5	0-14	D	D
	6	0-11	C	C
	7	0-21	B	B
	8	0-11	B	B
	9	0-13	C	C
	10	0-13	B	B

TABLE 2 (*continued*)

Day		Trials	Time (min -sec)	Doors chosen	Dark door
6	Morning	1	0-22	D	D
		2	0-12	B	B
		3	0-14	C	C
		4	0-12	B	B
		5	0-13	D	D
		6	0-16	C	C
		7	0-10	B	B
		8	0-21	D	D
		9	0-11	C	C
		10	0-9	B	B
	Afternoon	1	0-16	D	D
		2	0-11	B	B
		3	0-12	C	C
		4	0-10	B	B
		5	0-15	D	D
		6	0-11	C	C
		7	0-11	B	B
		8	0-12	D	D
		9	0-11	C	C
		10	0-11	B	B
7	Morning	1	0-18	D	D
		2	0-12	B	B
		3	0-13	C	C
		4	0-15	B	B
		5	0-14	D	D
		6	0-12	C	C
		7	0-11	B	B
		8	0-25	D	D
		9	0-10	C	C
		10	0-12	B	B
	Afternoon	1	0-10	D	D
		2	0-13	B	B
		3	0-13	C	C
		4	0-10	B	B
		5	0-12	D	D
		6	0-11	C	C
		7	0-13	B	B
		8	0-10	D	D
		9	0-13	C	C
		10	0-12	B	B
			All doors open		
8	Morning	1	0-10	D	D
		2	0-12	B	B
		3	0-10	A	A
		4	0-12	C	C

TABLE 2 (*continued*)

Day	Trials	Time (min-sec)	Doors chosen	Dark door
	5	0-14	D	D
	6	0-20	B	B
	7	0-12	D-A	A
	8	0-11	C	C
	9	0-12	D	D
	10	0-10	B	B
Afternoon	1	0-9	D	D
	2	0-9	B	B
	3	0-9	A	A
	4	0-12	C	C
	5	0-13	D	D
	6	0-8	B	B
	7	0-9	A	A
	8	0-10	C	C
	9	0-10	D	B
	10	0-9	B	D
9 Morning	1	0-29	D	D
	2	0-17	B	B
	3	0-10	A	A
	4	0-13	C	C
	5	0-12	D	D
	6	0-10	B	B
	7	0-11	A	A
	8	0-18	C	C
	9	0-10	D	D
	10	0-12	B	B
Afternoon	1	0-11	D	D
	2	0-12	B	B
	3	0-9	A	A
	4	0-14	C	C
	5	0-14	D	D
	6	0-12	B	B
	7	0-9	A	A
	8	0-13	C	C
	9	0-10	D	D
	10	0-9	B	B
10 Morning	1	0-47	D	D
	2	0-10	B	B
	3	0-10	A	A
	4	0-19	C	C
	5	0-12	D	D
	6	0-21	B	B
	7	0-8	A	A
	8	0-17	C	C
	9	0-15	D	D
	10	0-10	B	B

TABLE 2 (*continued*)

Day	Trials	Time (min.-sec.)	Doors chosen	Dark door
Afternoon	1	0-15	D	D
	2	0-19	B	B
	3	0-10	A	A
	4	0-16	C	C
	5	0-11	D	D
	6	0-10	B	B
	7	0-12	A	A
	8	0-11	C	C
	9	0-11	D	D
	10	0-12	B	B

A number after a letter in the column *Doors chosen* indicates the number of times the bird chose that door *To A* in the same column indicates an attempt to enter that alley even though it was blocked by the closed door

These results show 200 trials for mastery of the problem, with an average time of 20 25 seconds per trial and an average error of .386 per trial The time was reduced from the first day's average of 50 75 seconds to the last day's average of 12 09 seconds The average daily error was similarly reduced from 2 per trial to 0

The retention test, made after six days, showed no errors and an average time per trial of 11 33 seconds

TABLE 3
RESULTS OF THE TWO-CHOICE COLOR PROBLEM

Day	Trials	Time	Errors
1 Morning	1	0-32	2R
	2	0-27	2R
	3	0-55	3R
	4	0-26	1R
	5	0-26	1R
	6	0-20	1R
	7	0-9	
	8	0-10	
	9	0-8	
	10	0-10	
Afternoon	1	0-8	
	2	0-8	
	3	0-9	
	4	0-8	
	5	0-9	
	6	0-30	2R

TABLE 3 (*continued*)

Day	Trials	Time	Errors
	7	0-9	
	8	0-8	
	9	0-18	1R
	10	0-19	1R
2 Morning	1	0-8	
	2	0-9	
	3	0-10	
	4	0-9	
	5	0-13	
	6	0-9	
	7	0-8	
	8	0-9	
	9	0-9	
	10	0-8	
Afternoon	1	0-8	
	2	0-12	
	3	0-9	
	4	0-13	
	5	0-9	
	6	0-12	
	7	0-13	
	8	0-13	
	9	0-14	
	10	0-10	
3 Morning	1	0-7	
	2	0-10	
	3	0-10	
	4	0-9	
	5	0-10	
	6	0 9	
	7	0-10	
	8	0-9	
	9	0-10	
	10	0-9	
Afternoon	1	0-9	
	2	0-8	
	3	0-9	
	4	0-9	
	5	0-8	
	6	0-10	
	7	0-10	
	8	0-11	
	9	0-8	
	10	0-10	

The *R* in the error column indicates choice of the wrong light (red), and the number in front, the frequency of the error

These results show 30 trials for mastery of the problem, with an average time of 12.31 seconds per trial and an average error of .23 per trial. The average daily time was reduced from the first day's average of 17.45 seconds to the last day's average of 9.25, with a corresponding reduction in error of .7 per trial to 0.

Discussion

I think that part of the success in training this bird can be attributed to two factors: first, the bird showed no fear of the experimenter and was easily transferred by hand from cage to box, and, second, the bird was not starved or punished, as in the case of most such mastery tests, but was rewarded for correct choice. As the bird was never deprived of food, hunger could not have inhibited the learning.

The results of the two-choice light-and-dark problem show rather typical learning. The errors and the time decrease at about the same rate.

The learning in the four-choice box shows how easily a position habit can be formed and how difficult such a habit is to correct. After this position habit had been corrected, the learning progressed rather smoothly.

Undoubtedly, some of the speed in learning the color problem in the two-choice box was due to the fact that the bird was already quite familiar with the box and had only to learn to recognize the correct stimulus. Only errors of choice were made here, while in the first problem in this box several other errors were made, and the first day's average running-time here is slightly more than that of the last day's running-time in this box for the first problem. The fact that the bird was more positive to the red than to the green light in the preliminary tests may be due to a carry-over from the first two problems where the dark alley was the correct one, as the red is the less intense of the two colors.

University of Chicago
Chicago, Illinois

GIFTED CHILDREN AS HIGH-SCHOOL LEADERS

Frank H. Finch and Herbert A. Carroll

That intellectually gifted children lead scholastically has long been admitted; that they lead socially has long been denied. The prevalent belief is that the average student, admired by his fellows because of his lack of interest in things of the mind, holds most of the elective offices in school, in college—and in life.

The present writers, desiring to make a preliminary attack on the problem—admittedly a difficult one to solve—collected data bearing on the question from the records of the University High School at the University of

Minnesota The source of the data is an important fact to keep in mind University-high-school students are almost invariably superior children, and those at Minnesota are no exceptions. The mean intelligence quotient of the 211 pupils enrolled in the senior high school for the year 1930-1931 was 118.1 with a standard deviation of 12.2 and a range of 83-149

Method of Selection

The records of all pupils who were members of the senior classes from 1927 to 1932 were examined All such students who were found to have an IQ of 130 or above were placed in a group thereafter called "gifted," following Terman's terminology, those with IQ's ranging from 115 to 120 were placed in a group thereafter called superior (though they were actually average for their school); those with IQ's from 95 to 105 were placed in a group thereafter called average

Records from five group tests were available as a measure of intelligence Tests employed were Army Alpha 8, Pressey Senior Classification, Haggerty Delta 2, Terman Group Test, Form A, and Miller's Mental Ability Test, Form A Intelligence quotients were computed from the score of each test and equated by the method devised by Miller (1, p 360), so that the distribution of recorded intelligence quotients is similar to that which would have been obtained by use of the Stanford-Binet Of the five intelligence quotients of each individual, the median quotient was taken as the best measure of his ability

Pupils from these three intellectual levels were then matched on the factors of sex and number of years at the University High School No student was selected who had been in attendance less than three years, as "being acquainted" is obviously an important factor in leadership Matched cases were usually, but not always, in the same graduating class The members of each group are distributed rather evenly throughout the six years Furthermore, the changes in the extra-curricular program during this time have not been such as would give students at any given period an opportunity to receive a higher rating than at any other period

It was possible to match 66 pupils, in each of three groups, on this basis

Evaluating Leadership

The measure of leadership employed was the number of positions in extra-curricular activities to which an individual had been elected by his fellows No effort has been made to weight various positions Instead, one point has been assigned for each year in an elective position for which there is no scholarship requirement, and for which there is no faculty influence directly affecting the choice Offices and positions credited include the following

Senate (student government body) membership

Class officers
Boys' Club officers
Girls' Club officers
Managers and captains of major sports
Editors and business managers of publications

Records of positions indicating leadership were taken from the high school *annual, published by members of the senior class. In the few instances* where members of the senior class of 1932 were involved this information was obtained during the fall quarter by asking all members of this class for a list of positions held.

Analysis of Data

TABLE 1
LEADERSHIP RECORDS OF GIFTED, SUPERIOR, AND AVERAGE HIGH-SCHOOL PUPILS

	Gifted	Superior	Average
N	66	66	66
Years in High School	3.80	3.80	3.80
Intelligence:			
M	136.91 ± .455	117.83 ± .150	102.0 ± .243
σ	5.48	1.81	2.97
Chronological age in months			
M	157.95 ± .657	165.96 ± .653	173.13 ± .668
σ	7.92	7.87	8.05
Leadership points			
Total no.	120	103	60
Av. no.	1.82	1.56	0.91

Examination of Table 1 reveals two interesting facts: first, that the gifted excel both of the other groups in the number of leadership points earned, *amassing exactly twice as many as the average students*; *second*, that they did this even though they were eight months younger than the superior group and a little over fifteen months younger than the average group. Chronological age probably has considerable bearing upon high-school leadership, yet, in spite of this handicap, the gifted excelled.

When the three groups were divided on the basis of sex, the advantage still lay with the more intellectual, although the difference between gifted and superior girls is not so great as between gifted and superior boys. Average girls (in this school "dull" girls) appear to have but little chance to hold school offices (see Table 2).

TABLE 2

THE FACTOR OF SEX IN LEADERSHIP

	Gifted		Superior		Average	
	Boys	Girls	Boys	Girls	Boys	Girls
N	31	35	31	35	31	35
Years in High School	3.80	3.80	3.80	3.80	3.80	3.80
Intelligence						
M	137.70 ± .636	136.20 ± .625	117.50 ± .250	118.07 ± .173	101.89 ± .348	102.1 ± .340
σ	5.25	5.48	2.06	1.52	2.87	2.98
Leadership points (av.)	2.06	1.60	1.64	1.48	1.19	0.63

No attempt was made to find the reliability of the differences between the several leadership point averages. The distributions were so skewed that the probable errors would have meant little. The writers felt that a more satisfactory statistical evaluation could be made on a percentage basis (see Table 3).

Of the total number of leadership points earned by the three groups being studied, the gifted amassed the most. Although the difference between the percentages of the gifted and superior is not completely reliable—having a critical ratio of but a little over 2—it is indicatory of the probable trend. At least, it tends to show that the gifted child is not the unpopular little unfortunate that he is often pictured as being. Even with an N as small as 66, the difference between the superior and average, and between gifted and average is clear cut and reliable.

TABLE 3

THE RELATIONSHIP BETWEEN ABSTRACT INTELLIGENCE AND NUMBER OF LEADERSHIP POINTS

	Gifted	Superior	Average
N	66	66	66
Percentage of total number of leadership points	42.4 ± 1.98	36.4 ± 1.93	21.2 ± 1.64

DIFFERENCES BETWEEN LEADERSHIP POINT PERCENTAGES

Gifted—Superior = 6.0 ± 2.76
Superior—Average = 15.2 ± 2.53
Gifted—Average = 21.2 ± 2.57

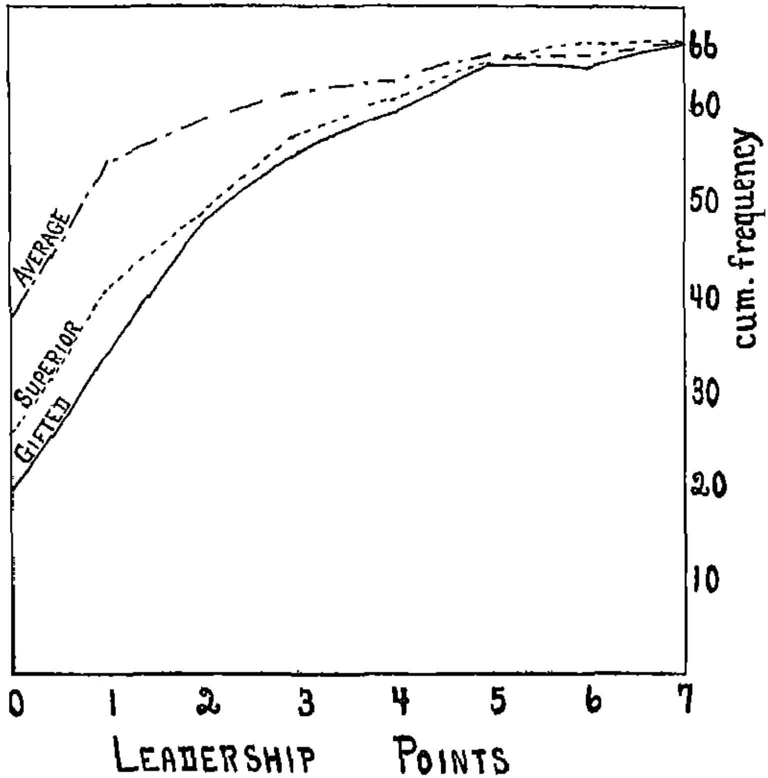


FIGURE 1

COMPARATIVE LEADERSHIP RECORDS OF GIFTED, SUPERIOR, AND AVERAGE HIGH-SCHOOL PUPILS

In Figure 1 the data have been presented graphically The significant fact emerging from this graph is the comparatively small number of gifted children who have not held at least one office

CONCLUSIONS

The number of children used in this study is too small to warrant any dogmatic statements The fact, too, that this school has a highly intellectual group of pupils enrolled may have affected the results It is quite possible that the gifted child in a typical public school would not manifest the same degree of superiority in the matter of leadership that he shows in the University High School in which the present data were gathered. Never-

theless, the figures presented in this brief discussion indicate rather definitely that, given a superior group of children to lead, the leading will tend to be done by the gifted children, while those who are at the lower end of the curve of the group have comparatively little opportunity to be elected to popular office These findings are in agreement with those of Terman, Hollingworth, Lamson, and Gray who have conducted more elaborate surveys

REFERENCES

1 MILLER, W S The variation and significance of intelligence quotients obtained from group tests *J Educ Psychol*, 1924, 15, 359-367.

University of Minnesota
Minneapolis, Minnesota

A NOTE ON THE LEARNING OF ELEVATED MAZES BY RATS

W D COMMINS

One of the most interesting observations made incidentally to a study of the effect of castration upon learning ability in rats was the comparative ease with which the animals mastered elevated mazes of the Miles type employed in the experiment One elevated maze was somewhat similar in pattern to the Warden U-maze, and consisted of 12 units The median number of trials required by a group of 257 animals to satisfy a criterion of three errorless trials out of four was 9 3 Although the rats were introduced to this instrument after they had learned the Stone multiple-T maze and the Stone multiple-light discrimination box, it hardly seems probable, in the light of other experimental data on transfer of training that they carried over sufficient increase in facility of adapting to a new learning instrument to account, on that basis alone, for their singularly better records on the elevated maze.

The ease with which the animals learned the first elevated maze made feasible the introduction of a second, after the first had been mastered. In this maze the sections were so joined that they formed a multiple-T elevated maze of 20 units Now, to our surprise, it was found that the rats satisfied the foregoing criterion of mastery with fewer trials than in the case of the first elevated maze They required but 4 9 trials as opposed to 9 3 for the first maze, despite the fact that the second might be expected, on the basis of the increased number of alleys and possible interfering habits, to offer greater difficulty than the first

These unexpected results from the two elevated mazes, along with observations of the animals' performances, stimulated the writer to formulate a hypothesis to explain the apparent anomalies It is the writer's opinion

that the greater ease with which rats master elevated mazes, as compared with floor mazes of similar type and number of alleys, results from the greater use they make of vision and visual cues. On the high relief maze the animals are up where a view of the neighboring units, as well as the general contour of the pattern, can be obtained, something that is not true for the alley maze with its walled-in pathways Miles (1) believed that rats trained upon the high relief mazes used visual cues in finding their way about His control of olfactory cues, by reversing and interchanging sections, seems adequate to rule out the effective use of that sense

In attributing the great facility with which rats learn the elevated mazes to their greater use of vision, one need not assume a high degree of visual acuity for the rat. Indeed, their behavior on the mazes seems to suggest relatively poor vision, for they very often ran down the blind alleys with such surety that they had great difficulty in avoiding a fall from the end of the run-way

The first trial on the elevated mazes was successful only after a rather long interval of time, but the facility with which the animals decreased their running time from then on would not be so surprising if we accord to the sense of vision an outstanding rôle in this adaptation The learning of the elevated mazes does not, apparently, so much involve the building up of specific response to particular visual cues (for this should conceivably take a longer time than our records show), as it implies the organization of a rather dimly perceived visual field, with emphasis upon the direction of the goal. To the writer, the rats' behavior on this type of instrument seems to border close upon, if it does not actually involve, the factor of insight The ability of the animals to perceive the elevated maze, as made up at least of four or five generous sweeping curves, instead of the twenty artificial sections of which it was constructed, would tend to make the instrument more simple than it might analytically appear. The use of vision would seem to be very well adapted to this type of organization, and we should not have to assume any great visual acuity to account for it. It is probably significant in this connection that Tolman and Honzik (2) obtained the most satisfactory evidence of insight, as it seemed to them, when their rats performed on the elevated type of maze.

References

1. Miles, W. R. The comparative learning of rats on elevated and alley mazes of the same pattern *J Comp. Psychol*, 1930, 10, 237-261
2. Tolman, E. C, & Honzik, C H "Insight" in rats *Univ Calif Publ. Psychol*, 1930, 4, 215-232.

Catholic University of America
Washington, D. C.

IDENTICAL SENSORY ELEMENTS VERSUS FUNCTIONAL EQUIVALENTS IN VISUAL TRANSFER REACTIONS[1]

By FRED McKINNEY

Lashley has given the name "equivalent stimuli" to those different stimuli which evoke the same reaction in an animal He (2, 3) has shown by the use of numerous patterns how a number of equivalent stimuli may exist for each stimulus used in training, and he has used this as further evidence against the theory of habits based on specificity of conduction paths which "demands that the reaction learned be given only to the excitation of the same receptor cells as were stimulated during training, or at least a significant number of them, these receptor cells being connected with the effectors by a system throughout which there is a point for point correspondence"

Thorndike, in his discussion of the general rationale of mental discipline says "Any assumption of gain in concentration, will-power, imagination, appreciation, conscience, reasoning or the like which cannot be described as a set of changes in the bonds between specified situations and definable responses, is extremely risky, and probably depends upon magic efficacy of mythical powers. Second, although every change must be in a specified bond, and though as a rule these bonds are between concrete, particular responses, some of these particularized bonds are of very widespread value" In another connection he states that " one or another element of the situation may be prepotent in determining the response" One holding such a viewpoint may claim that the equivalent stimuli contain certain identical elements of the original stimuli or involve "specific bonds," and it is these which are prepotent in evoking the particular response. However, Lashley (2) has shown the transfer of the habit of brightness discrimination from one eye to the other, which had been covered during training In this experiment it was impossible for the same sensory cells to function in both the formation and the performance of the habit The transfer took place readily in one animal without a visual cortex so that it is doubtful that one set of fibers was affected through central drainage by the set leading from the other eye But, since nothing is known about the actual central nervous processes involved, the matter is still open to some question

It is attempted in this experiment to gather data which will pertain to the question whether the response to equivalent stimuli is due to identical sensory elements or some sort of functional equivalence described by Lash-

[1]Experiment performed in the Psychological Laboratory of the University of Chicago and supported by a grant from the Otho S Sprague Memorial Institute The author is indebted to Dr K S Lashley for counsel and criticism

ley (2) as " the excitation of any cells of a system in certain ratios, . the response may be given to the ratio even though the particular cells involved have not previously been excited in the same way during the formation of the habitual reaction"

Six young pigmented rats were trained on a modified form of the apparatus described by Lashley (3) The apparatus differed from his in that the openings through which the rat jumped were 1½ inches apart instead of 2 inches and the pedestal stand from which the rat jumped was enclosed on the sides by two opaque boards leading to the screen and enclosed on top by a glass plate This prevented the animal from becoming disturbed by extraneous stimuli The same method of training was used as that described by Lashley The rats were rewarded with food, punished with a fall, and 20 errorless trials were required before the pattern was considered mastered After several months of training only two rats were successful in learning the rather difficult outlines, consequently, these rats were used in a number of discriminations described in this paper

The stimuli used were very similar to those which Lashley (3) used in his study of cerebral function in pattern vision They are shown in Figure 1. The aim in constructing such patterns was to provide stimuli which would have some resemblance in proportion and arrangement of parts but which, when projected from a constant distance upon the retinal surface would provide a minimal amount of indentical stimulation of retinal elements The devisal of such patterns offers difficulties, as the patterns must be simple enough to be readily distinguished by the rat and the difference between the patterns must be great enough to obviate an identity of the retinal images through defects of the refracting media of the eye

Lashley (5) has shown that sharp images of figures of this type are formed on the rat's retina under the conditions of this experiment and that acuity is great enough to assure that the images of patterns such as A and B will not be fused by optic imperfections This fact eliminates the possibilities of explaining any transfer of reaction on the basis of identical, vague patches spreading over the same area of the retina as the original stimulus

The following procedure was used in testing for equivalence The animals were first trained to choose one member of a pair of patterns, such as A A′ They were then confronted with a strange pair such as B B′ and given 20 "critical trials" with these In the critical trials both patterns were left unbolted so that the animal received food at every trial, that is, the spontaneous choice of the animal was recorded After each set of critical tests the animal was retrained on the original stimuli

In the first test with Rat 3 the method of comparison of rates of learning was tried The animal was first trained to choose stimulus X X′ This required 108 trials and 81 errors to reach the criterion of 20 consecu-

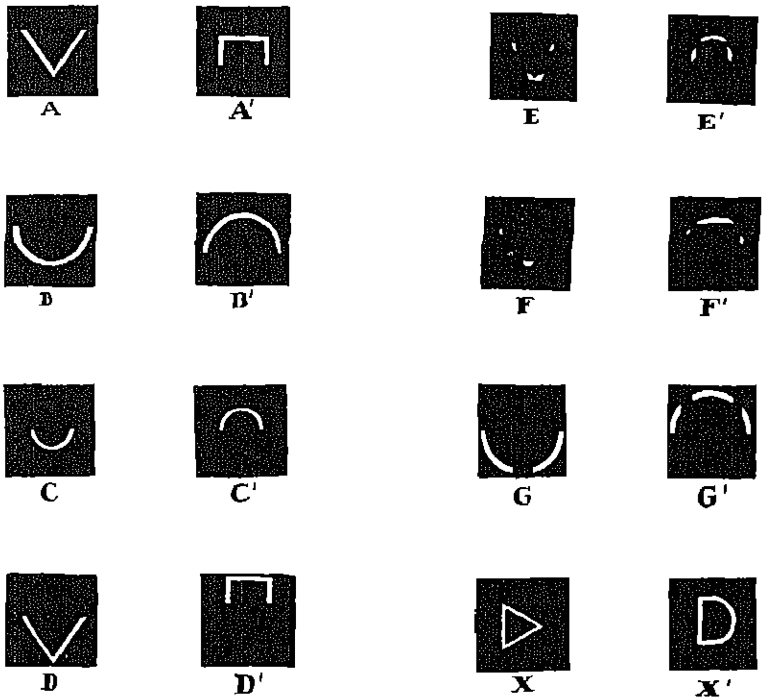


FIGURE 1
STIMULI TO WHICH THE RATS RESPONDED

tive errorless trials. The animal was then trained to choose A and avoid A′. This required 139 trials with 71 errors. The animal was next trained to choose B and avoid B′. This required 42 trials and 12 errors. There is evidence here for transfer from patterns A A′ to B B′ since the transition from X X′ to A A′ entailed no such reduction in training as that from A A′ to B B′. This method of using the saving in training as an index of transfer is less satisfactory than the method of critical trials, so that the latter has been used in other tests.

The records of critical tests and of retraining are presented in Table 1, in the order in which the tests were given. Rat 6, trained on A A′, required 129 trials and 73 errors to reach the criterion. Given 20 critical trials with B B′, she chose B 18 and B′ 2 times. With both animals and with two methods there is thus clear evidence of transfer of the effects of training from A A′ to B B′. It will be noted that there is a minimum of identical sensory elements in the two pairs of patterns.

To ascertain whether this transfer was due to the excitation of the few identical anatomical elements apparently still existing (common retinal cells and common synapses in the performance of two reactions) or to common relations without identity of elements in the figures, the following controls were used.

To learn whether the rat was reacting to the distance of the white pattern from the edge of the card or to the black area between the white pattern and the edge of the card rather than to the white figure on the card, the size of the circular patterns was reduced as in stimuli C C'. Both rats were trained again on A A' after $2\frac{1}{2}$ months had elapsed. (This time accounts for the poor records in the first retraining tests with A A'.) They were tested by the method of critical trials with C C', this time being retrained with A A' after 10 critical trials were given. This change in procedure was suggested for better results by some work done by another investigator in the laboratory. One rat responded to C 18 times out of 20, the other 20 out of 20, showing that the rat was reacting to the white figure as such rather than to some pattern formed by part of the figure and ground and common to A and B.

Both rats were again trained to A A' and, after they had met the criterion of learning, were subjected to patterns D D'. This control was inserted to make certain the animals were not responding to the position of white on the card, in which case the common element in A and B might be direction of gaze, a common motor element. In accomplishing this control the original positive stimulus was lowered in position, so that the white points which previously occupied a relatively high position on the card now occupied a relatively low one. The negative stimulus was raised, also reversing the positions of the two points of the figures. The rearrangement of the original stimuli on the card did not affect the response of the animals. Each of them responded to D 20 times out of 20 trials.

After again retraining on the original stimuli A A', stimuli E E' were presented to the animals. These figures represented the elements common to the original figures A A' and the equivalent stimuli C C', and were constructed by superimposing one figure on another with the aid of tissue paper and by marking the overlapping parts. One animal responded to E 16 times out of 20 critical trials and the other 20 out of 20. In a like fashion the elements common to A A' and stimuli B B' were constructed and are represented in stimuli F F'. One animal responded 19 out of 20 trials to F and the other 17 out of 20, as is shown in Table 1.

In these tests in which the identical physical parts of the stimuli were presented the animals gave evidence of transfer but less perfect than when the non-identical parts of B B' and C C' were also included. Furthermore, figures E E' and F F' obviously retain also some of the relational characteristics of A A' so that transfer to them is not crucial evidence of the importance of physical identity.

Finally, patterns were drawn which did not overlap any parts of A and A'. These patterns are labeled G G' and it is clear that they are designed so that they will not stimulate the same sensory cells excited by the patterns A A' and yet be similar to these patterns. Both rats responded to G 20 times out of the 20 trials, thus showing perfect transfer to a stimulus which apparently involved no common retinal elements with the original stimulus, but yet appeared similar in pattern.

How can these data be explained in terms of current theories of integration? The atomistic theory states that an overt activity is the result of the excitation of definite sensory cells attached to certain sensory nerve fibers leading to specific synaptic junctions in the central nervous system where the impulse affects other specific neurons. These neurons lead either directly or indirectly to effector organs. The crossing of the impulse at the synapse lowers the resistance offered by the junction, and thus, in the case of each overt activity, there is a lowered resistance at certain specific synapses. A single act may consist of many of these point-to-point connections, the activity being aroused when a sufficient number of them are stimulated. According to this theory the arousal of a perfected habitual act could only occur through the traversal of certain definite pathways which previously had been traveled and the synapses of which now offered lowered resistance. This theory would not grant that the pattern could fall on cells not stimulated and arouse fibers not traversed during the perfection of the habit, and evidence which might point to such an occurrence would be incompatible with the theory.

Do these data comply with such a theory? It has been shown that there was transfer to certain elements of the original situation. These elements, however, retained a certain amount of their relational characteristics. Also, the transfer in this case was less perfect than when the non-identical parts of B B' and C C' were also included. Further, the lack of transfer from X X' to A A' stimuli, which contained some common elements indicates that transfer is not just in terms of different physical elements common to the training stimulus and other diverse figures.

If common physical elements in the pattern are not the basis for transfer, there may be a basis in certain common motor responses. Hull (1) has shown that there are certain acts in a series of activities whose sole function is to serve as stimuli for other acts. He has called these pure stimulus acts. In this experiment the preliminary adjustment which the rat makes, his gaze, for example, may serve as a pure stimulus act. This factor, however, was controlled in this experiment by varying radically the relative positions of the stimuli on the cards. A change of this nature should disturb an animal using such a cue as a basis for its reaction, but neither rat was disturbed at all by this change.

The only other explanation of transfer is functional equivalence. The

TABLE 1

LEARNING AND TRANSFER RECORDS

Stimuli		Rat 3				Rat 6			
		Learning		Critical trials		Learning		Critical trials	
		Trials	Errors	Stimulus 1	Stimulus 2	Trials	Errors	Stimulus 1	Stimulus 2
A	A′	139	71			129	73		
B	B′			42	12*			18	2
A	A′	65	15			118	24		
C	C′			8	2			10	0
A	A′	10	3			23	1		
C	C′			10	0			10	0
A	A′	20	0			60	3		
D	D′			20	0			20	0
A	A′	29	3			20	0		
E	E′			10	0			10	0
A	A′	20	0			67	12		
E	E′			10	0			6	4
A	A′	20	0			50	4		
F	F′			9	1			7	3
A	A′	24	4			28	1		
F	F′			10	0			10	0
A	A′	20	0			20	0		
G	G′			20	0			20	0
X	X′	108	81						

*Trials 42, errors 12, by relearning method

data seem to support this explanation best, especially in view of the fact that there was perfect transfer to a stimulus G G′ in which there was no physical identity but only similarity in pattern.

It may be that students of psychology have always meant identity of relations involved when speaking of identical elements, but theoretical discussions concerning the nature of learning and transfer, particularly with respect to the doctrine of formal discipline show that many have had in mind identity of the structural elements mediating habits and this position is untenable

REFERENCES

1. HULL, C L Knowledge and purpose as habit mechanisms *Psychol. Rev*, 1930, 37, 511-525
2. LASHLEY, K S. Learning I Nervous mechanism in learning Chap 14 in *The foundations of experimental psychology*, ed. by C Murchison Worcester, Mass Clark Univ. Press, London Oxford Univ Press, 1929 Pp. 524-563
3. ————. The mechanism of vision I A method for rapid analysis of pattern vision in the rat. *J Genet Psychol*, 1930, 37, 453-460
4. ———— The mechanism of vision IV. The cerebral areas necessary for pattern vision in the rat *J Comp Neur* (in press)

5 ———— The mechanism of vision: V. The structure and image-forming power of the rat's eye *J. Comp Psychol.* (in press).
6 THORNDIKE, E L Psychology of learning. New York: Teach. Coll Columbia Univ, 1913

University of Missouri
Columbia, Missouri

A FURTHER NOTE ON SIZE PREFERENCE OF ALBINO RATS[1]

JOSEPH G YOSHIOKA

In the writer's note (3) on size preference of albino rats a remark was made as to a bad judgment shown by the rats in preferring large sunflower seeds to small ones because large seeds required more time and labor to eat but did not contain any more meat than small ones Professor Tolman (2, p 31) expressed his opinion that this preference would appear to be a mere "fixation," not a judgmental response, unless it could be shown that the rats would show "docility" in changing their preference to one for the seeds, which, in the long run, would satisfy their biological needs better. To demonstrate such "docility," if possible, further tests were given to the same rats

It is to be recalled that in the first test, lasting ten days, 21 rats ate daily 12 13 large sunflower seeds on the average, as against 10 59 small seeds in *five minutes The preference for large seeds, however, was better shown in* the first five days than in the last five days. Table 1 testifies to this statement

TABLE 1
AVERAGE SEEDS EATEN DAILY
N = 21

	Large	Small	Difference	Critical ratio
First five days	13.15 ± 45	10 82 ± 47	2 29	3 52
Last five days	11 28 ± 49	9 94 ± .60	1 34	1 73

This decrease in the degree of preference for large seeds may have been due either to their becoming somewhat indifferent to the choice toward the end or to their gradual learning that small seeds satisfied them better The writer inclined at that time to sponsor the first assumption from experiences in maze running which decreases in accuracy when a test is unduly prolonged Since the object of the experiment was to demonstrate size preference, the subsequent tests were given for five days only

[1]Studies from the Behavior Research Fund, Chicago Series B, No. 186

During the next five days the rats were tested daily in darkness, and it was found that the preference disappeared. For the next ten days the animals were given large or small seeds exclusively on alternate days to measure average speed to eat each kind of seeds. Small seeds were found to be eaten with greater speed and ease. Next, five tests were given on every other day. On the test days they were fed twice as much as normal after the test, but on the intervening days they were not fed at all. Under intensified hunger thus induced they ate 10.64 large seeds, on the average, as against 9.79 small seeds. The critical ratio between the two means was 1.31. Preference for large seeds reappeared, but, as compared with that shown in the first test, it was not so marked, where the critical ratio was 3.52 in the first five days and 1.73 in the last five days. Next five tests were given daily immediately after normal feeding. The average number of large seeds eaten in the tests was 11.12 and that of small seeds 11.54. The difference between the two means was not significant, the critical ratio being .43. The preference disappeared, but it is to be noted that here small seeds were eaten slightly more (although not significantly) than large ones, and both seeds were eaten more than in the preceding tests under hunger. The latter incident was explained by the observation that the rats in hunger made too much fuss in eating the seeds, splashing them about, picking them up in a hurry, and letting them slip out of their hands, while after a meal they sat quietly and methodically picked them up one after another without failure. Perhaps they were more deliberate in the choice, too, and selected small seeds more because they had begun to appreciate how good these were.

It was thought that more satiation might induce still more deliberate choice. Hence 20 rats (one had died) were fed twice as much as normal and tested for the next five days as before. Each day a retest was given also immediately after feeding. It was assumed that they were less hungry in the tests and still more so in the retests. The results are shown in Table 2.

It is seen that preferential choice disappeared in both pre- and post-feeding tests. Too much satiation seemed not inducive to preferential choice. It is to be noted, however, that in the pre-feeding tests large seeds were eaten, or rather tended to be eaten, slightly in excess of small ones, but in the post-

TABLE 2

Average Seeds Eaten Daily under Overfeeding

$N = 20$

	Large	Small	Difference	Critical ratio
Before meal	11.48 ± .60	11.31 ± .38	.17	.24
After meal	12.64 ± .59	12.59 ± .41	.05	.07

feeding tests any such tendency disappeared. These results were somewhat similar to those obtained previously under hunger and after a meal, only in a lesser degree.

Next, in order to induce better deliberate choice the animals were put back to normal feeding, and tested daily for five days in pairs with the seeds doubled in number. It is to be remembered that so far they had been tested individually with 20 large and 20 small seeds. Now 40 large and 40 small seeds were offered to each pair of rats. The pair faced each other across a pile of seeds, occupying a circular space, about four inches across. The pile kept the two rats sufficiently far apart so that no fighting between them was observed. Furthermore, to lessen rivalry the partners were selected from cage-mates. In fact, so many seeds were in front of one that the presence of the other on the opposite side was apparently ignored. The average number of large and small seeds eaten daily by these pairs of rats are given in Table 3.

It is seen that small seeds were eaten slightly in excess of large seeds. As compared with a similar preference shown once before, the degree of preference here was greater. In the post-feeding tests under overfeeding the animals chose small seeds slightly more than large seeds so that the critical ratio between the two means of large and small seeds was .43 equivalent to probability of .67, that is, the obtained difference may occur by chance 67 times out of 100. Here the obtained difference may occur by chance 40 times out of 100. Unless probability is .1 or less, the preference shown for small seeds cannot be said to be positively demonstrated, but in the comparison just made a tendency for such preference was more positively shown in the last tests than previously.

In the next daily test for five days, rivalry was made to play its part in

TABLE 3

AVERAGE SEEDS EATEN DAILY BY PAIRS OF RATS

N = 10 pairs

Large	Small	Difference	s	t	P
21.78	23.26	—1.48	4.009	.828	.4

$$s = \frac{1}{n_1 + n_2} \left\{ \Sigma\,(x_1 - m_1)^2 + \Sigma\,(x_2 - m_2)^2 \right\}$$

$$t = \frac{m_1 - m_2}{s} \sqrt{\frac{(n_1 + 1)\,(n_2 + 1)}{n_1 + n_2 + 2}}$$

P = probability of the obtained difference occurring by chance in 100.

N. B. Since the number of cases was halved by pairing, Fisher's method was used for group comparison. For the original formulas, see Fisher (1, p. 107).

TABLE 4
AVERAGE SEEDS EATEN DAILY BY PAIRS OF RATS UNDER RIVALRY
N = 10 pairs

Large	Small	Difference	*s*	*t*	*P*
12 32	12 14	.14	2,916	138	.9

the choice by halving the number of seeds and time for eating Now each pair was given 20 large and 20 small seeds and was taken out of the test enclosure at the end of two and a half minutes. The result is shown in Table 4

It is seen that preference disappeared almost completely The obtained difference may occur by chance 90 times out of 100 Under the stress of rivalry the animals helped themselves to whatever seeds were handy, disregarding their size or content Perhaps it was too much to expect that they would show deliberate choice under the conditions At any rate it was shown that social influence is not negligible in their behavior

It is concluded that (1) the albino rat prefers large sunflower seeds to small ones for apparent large content in the former (2) Such preference disappears under satiation when food is less demanded or under conditions of rivalry when quick rather than discriminate choice is demanded. (3) Preference for large seeds can be replaced by preference for small ones when the animal is tested under certain conditions where deliberate choice is made possible, and after he had been given sufficiently long experience to get acquainted with better quality of small seeds. (4) Hence size preference shown by the rat is not a "fixation" but a judgmental response

REFERENCES

1 FISHER, R A Statistical methods for research workers. London Oliver & Boyd; New York Stechert, 1928. Pp xi+269
2 TOLMAN, E C Purposive behavior in animals and men New York: Century, 1932. Pp xiv+463.
3. YOSHIOKA, J G. Size preference of albino rats. *J Genet. Psychol*, 1930, 37, 427-430

Anthropoid Experiment Station of Yale University
Orange Park, Florida

REGARDING THE LAW OF PARSIMONY

JOSEPH W. NAGGE

Lloyd Morgan, in his writings on animal behavior, has laid down a canon of interpretation which has come to be known to psychologists as the law of parsimony Morgan, in his *Introduction to Comparative Psychology*, has formulated the principle as follows "In no case may we interpret an action

as the outcome of the exercise of a higher psychical faculty, if it can be interpreted as the outcome of the exercise of one which stands lower on the psychological scale" (p 53). Thus Morgan would prefer to explain an animal's act in terms of perception rather than by postulating an ideational exercise of cognition, unless such an explanation were inadequate However, as the author explains, this canon of interpretation by no means prohibits the explanation of an animal act in terms of the higher mental processes, provided that there is at hand independent evidence of their occurrence in the agent

This canon seems to have undergone a transformation in general psychological usage until it might now be tentatively expressed thus. of any possible number of explanations of an animal act the simplest possible explanation should be employed.

The law of parsimony has been frequently invoked by psychologists in judging the merits of various explanations of psychological phenomena, particularly in the field of animal psychology Some of the more mechanically-minded behaviorists have utilized the principle against anthropomorphic interpretations of animal acts in terms of conscious processes Those so criticized have retorted that the principle of parsimony supports their explanations, and point out that we must attribute to the lower organisms conscious processes and purposes like our own, since our consciousness is the only one we intimately know, and hence the simplest of its kind.

We see, then, that the law of parsimony as generally conceived is itself subject to interpretation Is the simplest psychological explanation to be in terms of psyche, in terms of physiology, or in purely physical terms? What is the simplest explanation of an animal running the maze? The theologian, the chemist, the psychologist might all explain the action in varying terms. One psychologist would point to conditional reflexes as the simplest explanation, another might indicate closure or insight as the basic explanation, another might insist upon instinctive urge or conscious purpose. Would our principle of parsimony allow us to claim that explanations in terms of simple reflexes, or simple sensations, are more elementary and hence more satisfactory than explanations in terms of more complex *Gestalten* or conscious purposes?

Further, it is difficult to differentiate qualitatively between explanation and description in psychology Good authority has maintained that explanation is only more elaborate description Does this mean that the principle of parsimony should also apply in psychological description as well as in psychological explanation? Would the simplest description thus be the best? What effect would such considerations have upon the descriptive units of some of our psychological schools?

The principle of parsimony as a guide for psychological theorizing has not been an unmixed blessing While it has tended to sober our theorizing, it

has also served as a sanctuary for those who have formulated an over-simplified ideal in psychological explanation. Frequently we are led to wonder if our explanations should not be far more intricate and involved than we now generally view them Psychologists of long standing cannot help recalling Schopenhauer's position, along with others, when he indicated the human mind as the supreme enigma

If the principle of parsimony is to be utilized further in psychological theorizing it should be more sharply defined and delineated. At present it involves a troublesome phrase, "adequate explanation" What is an adequate explanation? Many feel that psychology as science has a legitimate right to its own psychological explanations A physiological explanation of psychological acts is also a possibility Which of the two is the simpler explanation? Before a principle of parsimony can become unambiguous for all explanations of psychological phenomena we must agree upon a definition and ideals for the science.

University of Chicago
Chicago, Illinois

BOOKS

EDITED BY

EDMUND S. CONKLIN
SANTE DE SANCTIS
A. R LURIA
GODFREY THOMSON
C. J WARDEN

S ZUCKERMAN *The Social Life of Monkeys and Apes* (International Library of Psychology, Philosophy and Scientific Methods) New York Harcourt, Brace, 1932 Pp. xii + 375 + 24 plates

In *The Social Life of Monkeys and Apes* Dr. Zuckerman, Anatomist to the Zoological Society of London and Demonstrator of Anatomy, University College, London, has produced an extremely interesting and highly valuable contribution to comparative sociology, psychology, and physiology Primarily engaged in studying the social behavior of sub-human primates, he approaches the subject, as he states in the introduction, "from the deterministic point of view of the physiologist, treating overt behaviour as the result or expression of physiological events which have been made obvious through experimental analysis" He reviews the literature, both historical and modern, including essential contributions from the fields of physiology of reproduction, comparative psychology, anthropology, and mammalian sociology To this mass of material, critically evaluated, he adds original observations made upon a colony of captive hamadryas baboons in London and wild chacma baboons in South Africa The subjects covered by the book can be inferred from the chapter headings: (1) Human Sociology and the Sub-Human Primates, (2) Mammalian Sociology; (3) Sexual Periodicity, (4) The Breeding Season and Society, (5) The Physiology of the Reproductive Processes of Mammals Other Than Primates, (6) The General Character of the Menstrual Cycle, (7) The Morphology, Physiology, and Interpretation of the Menstrual Cycle, (8) The Oestrous Cycle and Behaviour, (9) The Menstrual Cycle and Behaviour; (10) The Individual Within the Group, (11) Apes and Monkeys in the Wild General, (12) Baboons in South Africa, (13) The Social Groups of Wild Apes and Monkeys, (14) The Hamadryas Baboon Colony, (15) Dominance and the Liberation of Sexual Responses, (16) The Communal Life of the Baboon, (17) The Developments of Social and Sexual Responses; (18) Altruism and Society, (19) The Social Scale

In the early chapters the author criticizes and rejects all teleological

speculations upon, and explanations of, mammalian society, stating that "the chief subject matter of a scientific mammalian sociology is seen to be ecology, reproductive physiology, and those influences which can be classed together as due to the variations of the individual." The customarily drawn distinction between animal associations and animal societies in his estimation "does not seem to depend upon more than a superficial analysis of the facts of animal behaviour" He substitutes a classification of mammals based upon differences in their reproductive physiology which largely determine the nature and permanence of their social organization In this classification the three main groups include (1) primates, in which, according to the author, the female is ready to accept the sexual advances of the male at all times, and which live in family or big social groups, (2) non-primates or lower mammals that have an anoestrus, a demarcated mating season, and in which the social group commonly endures as long as it members are sexually potent; and (3) the remaining lower mammals which, like the primates, breed throughout the year, do not experience anoestrous quiescence of sexual activity, and which usually live alone or in small parties

In his chapter on "Altruism and Society" the author, still critical of the teleological interpretations of other writers, indulges in some interesting speculations which are less objective in nature than would be expected after reading the preface After considering the various social relationships and adjustments between living and between living and dead monkeys and apes, the author concludes the chapter with the statement that "Monkeys and apes do not recognize death, for they react to their dead companions as if the latter were alive but passive Thus the 'blind' character of their social responses would seem to deny them any claim to altruism" (p 305).

Some of the interesting conclusions offered by Dr Zuckerman concerning the social life of sub-human primates are as follows The overt socio-sexual activities of these animals are much farther removed from those of the lower mammal than from those of man The herd, the largest social unit of sub-human primates, may be formed from several family parties or harems, but never appears to be as stable a unit as the family The family party, nucleus of monkey and ape societies, consists of an overlord and his harem, held together primarily by the interest of the male in his females and by their interest in their young Organization within groups depends upon individual dominance The social level of the monkeys and apes is one through which man, also an Old World primate, may have passed in the pre-human stages of his evolution, prior to development of the monogamous family unit which, "according to most authorities," is typical of the lowest level of human society The change from polygyny to monogamy may have been caused by differences in diet for, "primitive man, who . . . was an animal largely dependent upon a diet of meat, would not have gone hunting if in his absence his

females were abducted by his fellows. Reason may have forced the compromise of monogamy—*If*," as he states elsewhere, "reason played a part . . ."

The principal criticism which can be made of the book is the author's assumption that the socio-sexual relations of apes are the same as those of Old World monkeys While this may be the case, there is still insufficient evidence available concerning apes, particularly in the wild state, to justify such a conclusion. The book's value remains unchanged, however, regardless of the primate level to which the conclusions apply. It is a new and interesting approach to the study of primate sociology. In addition, it is a storehouse of information concerning comparative reproduction and social behavior.

The book contains excellent photographs of captive hamadryas baboons, a classified index of species mentioned and a bibliography of 402 titles.

O L TINKLEPAUGH

BOOKS RECEIVED

EMERSON, H [Ed] *Alcohol and man: the effects of alcohol on man in health and disease* New York Macmillan, 1932 Pp. xii+ 451 $3 50.

FISHER, D C, & GRUENBERG, S. M [Eds] *Our children a handbook for parents* New York. Viking, 1932 Pp x+248 $2.75

JACOBSEN, C F, JACOBSEN, M M, & YOSHIOKA, J G Development of an infant chimpanzee during her first year *Comp Psychol Monog*, 1932, 9, No 1 (Serial No 41). Pp 94 $1 50.

JACOBSON, P B. Two experiments with work-type reading exercises in the ninth grade *Univ Iowa Stud Stud Educ.,* 1933, 8, No 5 (New Ser, *No. 247) Pp 85 $75*

JOHNSON, B J Child psychology Springfield, Ill., & Baltimore, Md. Thomas, 1932 Pp. xii+440 $4 00.

LULL, H G Secondary education Orientation and program New York Norton, 1932. Pp xviii+366

OGDEN, R M Psychology and education (Rev. in collaboration with Frank S Freeman) New York Harcourt, Brace, 1932 Pp xiv+350. $2 50

MURPHY, G, & JENSEN, F Approaches to personality: some contemporary conceptions used in *psychology and psychiatry* (With a supplement by John Levy) New York: Coward-McCann, Inc., 1932. Pp xvi+427 $3 75.

MURSELL, J The psychology of secondary school teaching New York: Norton, 1932 Pp xii+468.

TOLMAN, E C Purposive behavior in animals and men New York & London Century, 1932 Pp xiv+463 $5 00

Annual report of the board of regents of the Smithsonian *Institution*, showing the operation, expenditures, and condition of the Institution for the *year ending* June 30, 1931 Washington Govt. Printing Off, 1932 Pp xiv+592.

SUBJECT INDEX

Activity of preschool children, rating technique, 330
Adopted children in nature-nurture research, 294
Age differences in toy-preference of children, 417
Animal, bird, learning of, 463
  cockroach, effect of inactivity on learning and retention, 253
  kittens and simple alternation problem, 393
  Petz's conure, learning of, 463
  rat, association span of, 57
    conditioned response in, 91
    food preferences of, 430
    initial maze behavior, 358
    isolated action of compound stimuli in, 127
    learning of elevated mazes by, 481
    learning versus skill in, 406
    relative strength of primary drives, 16
    size preference of, 489
    social facilitation in, 211
Apparatus, life-size alley maze for children, 235
Association span in white rat, 57

Child, inventive genius of, 369
Children, adopted, in nature-nurture research, 294
  factors influencing sleep of, 222
  gifted, as high-school leaders, 476
  habits of, 167
  imaginary playmates of, 380
  learning and configurational factor, 3
    and maturation in, 36
  life-size maze for, 235
  ratings of, personal equation in, 267
  reliable IQ determination, 152
  school problems and mental hygiene, 439
  sidewalk drawings of, 192
  study of rating of activity, 330
  toy-preferences of, 417
Conditioned response in rat, 91
Configural factor in children's learning, 3

Drawings, sidewalk, Gestalt principles of, 192
Drives in rat, 16

Feeding of rats, social facilitation in, 211
Food preferences of rat, 430

Genius, inventive, in child, 369
Gestalt and children's learning, 3
  principles in sidewalk drawings of children, 192
Gifted children as high-school leaders, 476

Habits of nursery-school children, 167

Identical sensory elements *vs.* functional equivalents in visual transfer patterns, 483
Imaginary playmate, questionnaire study of, 380
Inactivity, effect on learning and retention, 253
Intelligence quotient, interpretation of, 116
    reliable, for young children, 152
  test scores and motivation, 136
Interpretation of IQ, 116
Inventive genius of child, 369

Kittens and simple alternation problem, 393

Law of parsimony, 492
Leaders, gifted children as, 476
Learning, configural factor in, 3
  kittens on simple alternation problem, 393
  and maturation in children, 36
  of Petz's conure, 463
  and retention in cockroach, 253
  vs. skill in rats, 406

Maturation and learning in children, 36
Maze behavior, rat, initial, 358
  for children, 235
Mazes, elevated, rat's learning of, 481

Mechanical ability, criticism of recent work, 228
Mental hygiene applied to children's school problems, 439
 status of adopted children, 294
Motivation and intelligence test scores, 136

Nature-nurture problem, factor of selective adoption in research, 294

Parsimony, law of, 492
Personal equation in ratings, 267
Petz's conure, learning of, 463
Playmate, imaginary, questionnaire study of, 380

Rat, association span of, 57
 conditioned response in, 91
 food preferences of, 430
 initial maze behavior, 358
 isolated action of compound stimuli in, 127
 learning of elevated mazes, 481
  vs. skill in, 406
 relative strength of primary drives in, 16
 size preference of, 489
Rating children, personal equation in, 267
 scales, personal equation in, 267
 technique in activity of preschool children, 330
Retroactive inhibition and inactivity, 253

Sex differences in toy-preferences of young children, 417
Size preference of rats, 489
Sleep of children, factors influencing, 222
Social facilitation in feeding of rats, 211

Transfer and identical elements, 483